Encyclopedia of Nanoscience and Nanotechnology

Volume 7 Number 1 2004

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanomechanical Properties by Nanoindentation

Carlos M. Lepienski

Universidade Federal do Paraná, Curitiba-PR, Brazil

Carlos E. Foerster

Universidade Estadual de Ponta Grossa, Ponta Grossa-PR, Brazil

CONTENTS

1. INTRODUCTION

Load and depth-sensing indentation, also known as the nanoindentation technique, started to be developed in the 1980s, and nowadays is being continuously improved. The experimental device is based on recording load, depth, and time during the indentation process. The physical models and experimental systems are based on contact mechanics adapted to conditions of very low applied loads and indenter displacements. This means sensitivity of the applied load and tip penetration on the order of micronewtons and nanometers, respectively.

The determination of the mechanical properties of nanostructured materials is a new and very interesting area. Nanostructured materials are characterized by nanoscale structures with a significant volume comprised of surfaces, interfaces, and grain boundaries, giving properties vastly different and often greatly superior to their bulk-metallic, ceramic, and polymeric counterparts. They can be produced by different techniques to form a myriad of engineered structures. These advances will provide better sensors, medical diagnostics, displays, data storage technologies, and surface protection coverings. However, the determination of their mechanical properties and how these properties are measured in nanostructured materials are in constant revision. As a result, an intense effort has been made in recent years to adapt and develop methods in order to measure deformation and fracture properties at the nanoscale. At the same time, experimental and theoretical work is being done to determine and understand the properties and processes influencing the mechanical behavior and fracture in small volumes. Advanced and novel characterization and testing methods, as well as analytical, continuum, and atomistic simulations, are being continuously developed. Of particular interest are those studies that extend over length scales from atomistic to nanometer or from nanometer to submicron, and thus provide insight regarding length-scale effects.

The characterization of mechanical properties at nanoscale includes: structure–property relationships at the nanoscale in nanostructured materials, composites, films, multilayers, and functionally graded materials; mechanical properties of bulk materials at the nanoscale; deformation and fracture processes at the submicron scale; dislocation emission at probes and crack tips at the nanoscale; effects of intrinsic stresses on properties and fracture; creep measurement and mechanism determination; tribological properties of nanostructured material; comparative results for the modeling of submicron-scale indentation, scratch, and wear response; and development of standards tests to measure mechanical properties at the submicron scale.

The purpose of this review is to show the reader the basic concepts of nanoindentation, and how it can be used to determine mechanical surface properties in the nanoscale regime. These concepts include the basic equations, a brief discussion about theoretical models, the equipment calibrations, and the correct choice of an appropriate indenter. In the following, focus will be given to a series of application results for modified surfaces and thin films (coatings),

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (1–20)

tailoring the use of this technique. The further tendencies about using load and depth-sensing indentation are also shown.

This chapter consists of four parts. First, a definition of nanoindentation and some comments with respect to how this technique can be useful for nanoscience and nanotechnology are given. Historical aspects, fundamental equations, and experimental methodology follow. A third part concerns the most important applications and principal results that can be found in the literature. Finally, some discussions about new applications, recent developments in the technique and in the theoretical approach, as well as future perspectives are also taken into consideration.

Nanoindentation is a term used to designate indentation experiments where very low loads are used to press a hard indenter into a surface in a controlled way at submicronmeter penetration depths. As an important feature of the process, the applied load, the indenter displacement, and the time are continuously recorded during the experiment. Due to its versatility, nanoindentation is one of the most important methods to obtain the mechanical properties in the near-surface regions, and it can be applied to practically all types of solid materials.

The Vickers and Brinnel conventional tests use the indentation image to measure the contact area which is necessary to calculate hardness. Unlike this, in nanoindentation, the contact area is obtained from the load and displacement data recorded in a complete load–unload cycle, without the necessity of imaging methods. If more elaborate analyses are performed, other mechanical properties, such as elastic modulus, residual stresses at the surface, strain rate sensitivity, toughness of brittle materials, and adhesion quality of thin films, can be determined by applying the nanoindentation technique.

The determination of nanomechanical properties is a new field, as well as practically all areas in nanoscience and nanotechnology. For this reason, there is a considerable gap to link the mechanical properties from a continuum mechanics approach to an atomistic point of view. This atomistic approach takes into account the forces and movements of individual atoms or groups of atoms that are not considered in continuous mechanics. Atomic force microscopy (AFM) gives us a better understanding of the weak mechanical forces that govern the behavior of atoms in the surface of the materials. However, AFM is not well suited to determine the behavior of the material in a region deeper than a few atomic layers under the surface. Nanoindentation that presents high depth and load resolution is very useful to correlate the nanoscale mechanical behavior with the material structure. In the near future, an improved modeling and interpretation of load versus displacement data is expected since a considerable effort is being developed in the nanoindentation area. The new incoming equipment will probably become better. For these reasons, nanoindentation will continue to play an important role in nanotechnology development.

Surface mechanical properties of several different materials, like metal nanocomposites, polymer films, metal surfaces, coatings, and microelectromechanical systems (MEMS), are being determined by applying the nanoindentation technique. Tailoring modified surfaces to improve mechanical and tribological applications requires a good knowledge of the mechanical properties at the material surface. Knowledge of the hardness, elastic modulus, and wear resistance is very important to specify the correct application for newly developed materials used in the electronic, automotive, or aerospace industry. Then, nanoindentation has an important role in the analysis of nanoscale mechanical properties of materials, and its application is continuously increasing. Finally, despite the fact that results will be more consistent, a great effort must yet be made to adequately interpret the depth and load-sensing indentation data because the deformation process under the indenter is very complex.

2. FUNDAMENTALS OF NANOINDENTATION

In this section, the basic principles of nanoindentation are presented. Historical aspects are described, and it is shown how this technique is used to obtain the surface mechanical properties of materials. Far from being a complete treatise about the theme, the present section can be considered more like a straight guide to understand how nanoindentation can be applied to the study of nanoscale mechanical properties of surfaces and thin films. More advanced works and extended reviews about this theme are indicated in the text.

2.1. Historical

Indentation experiments have been extensively used since the middle of the last century to measure the hardness of materials [1]. However, techniques for probing the mechanical properties of materials at nanoscale were developed only in the early 1980s. This mechanical probing consists of a load and depth-sensing indentation, which gives information about the surface mechanical properties of materials at depths lower than 1 μm. Pioneer works of Pethica et al. [2, 3] showed the first results of indentation at nanoscale that were obtained with an instrumental device that is known now as nanoindentation equipment. They developed an electromechanical apparatus that was used to obtain the hardness at nanoscale. A few years after that, a sequence of other articles were published, showing a great instrumental development, and a new research area on mechanical surface properties was established. The original equipment's purpose was to measure hardness at nanoscale. However, nanoindentation is now used in a wide spectrum, and the recorded data obtained from the indentation are applied to determine different mechanical properties. Other independent works were done in the same period [4, 5], and equipment of different designs, but for the same purpose, were constructed. All of these works were important for developing this new research area.

In conventional indentation methods such as Vickers and Brinnel tests, the hardness is obtained from the ratio of the applied load to the area of the residual impression, which is determined by image methods. However, for indentation at

nanoscale, imaging methods demand a long time, and very special techniques are necessary to obtain the actual area. Sometimes it is very difficult, if not impossible, to make good images from indentations when very low loads are used. Now, with the depth-sensing indentation, it is possible to measure the hardness without imaging the indentation impression.

Pethica et al. [2, 3] reported the use of an area function to calculate the indentation area for a determined indenter penetration depth. The method is based on the knowledge of the shape of the indenter, and on the fact that the material normally conforms to the shape of the hard indenter. The contact area was obtained from final depth after the unloading. However, this first method was not correct as it did not consider the elastic recovery of the sample.

Doerner and Nix [6] later elaborated a method to determine the hardness, and additionally the elastic modulus, from load versus displacement data. In their method, the unloading curve is considered to be governed by the elastic properties only. In addition, the elastic recovery is considered similar to that obtained from a flat cylindrical punch, where the contact area remains constant during the unloading process. Since the elastic recovery from unloading a flat punch is linear [7], they extrapolated the initial part of the unloading curve by a linear fit to obtain the contact depth, which gives better results than at maximum depth or at the residual depth that was previously used in the area function. As an improvement, with this new method, the elastic modulus of materials surface could then be obtained from the loading and unloading data. However, this approach does not give accurate results since the areas measured by imaging methods give different values when compared to the ones obtained by the Doerner and Nix [6] method.

Oliver and Pharr [8], in 1992, described the necessary approach to make nanoindentation a widely acceptable probing method to obtain hardness and the elastic modulus in practically every kind of surface. Their article is the most cited in the nanoindentation area at the present moment. The principle, called the Oliver and Pharr method, is practically a standard when it is necessary to obtain hardness and the elastic modulus from load–unloading curves using load and depth-sensing indentation. After that, a great number of papers were presented showing new improvements on how to obtain new properties of materials from nanoindentation data, and they are described in the following as each aspect is being discussed. In what follows, the basic principles are presented, and some recent improvements in analysis and data interpretation are discussed.

2.2. Basic Principles

2.2.1. Depth and Load-Sensing Equipment

A schematic diagram of equipment that is often used in depth-sensing indentation is shown in Figure 1. It consists of a system of a vertical axis supported by springs to a cell. The indenter is at the end of the axis, as shown in Figure 1. The system is composed of a force actuator, normally an electromagnetic shaker actuator, and a sensor of depth that is generally a capacitance displacement gauge. Other systems that use electrical force applied by capacitance plates

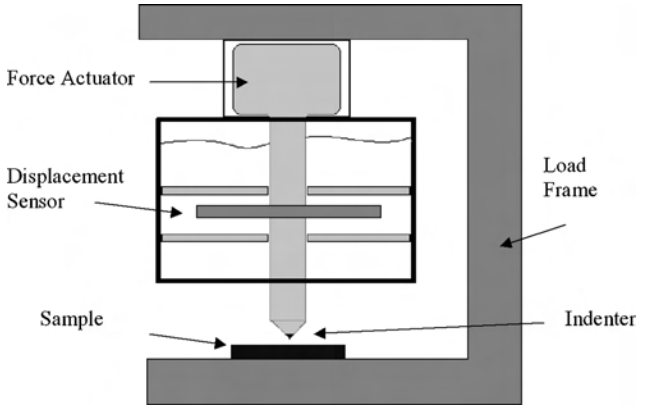


Figure 1. Schematic of a nanoindenter showing the various equipment parts.

or piezoelectric actuators are also used in different equipment designs.

The force actuator normally is capable of applying forces as lower as 1 μN, and the displacement gauge sensor can give a depth resolution better then 0.1 nm. However, for measurements at depths lower than 20 nm, additional rigor and care are needed to obtain useful results. The maximum load used in this kind of equipment is normally about 500 mN. On the other hand, if better displacement and load resolutions are necessary, the maximum load is normally lower, and the depth resolution is increased for shallow penetration. An extensive description with details of different equipments used in nanoindentation, as well as their calibration processes, is presented by Bhushan [9].

Some nanoindentation machines allow the user to build an indentation pattern in a two-dimensional array by using a previous programming. These arrays can consist of a combination of a linear x and y pattern in order to form different two-dimensional geometrical figures like a square, a rectangular, or also triangular shapes. A typical example of an indentation shape produced by a Nanoindenter XP™ is shown in Figure 2. The distance between each indentation can also be programmed, and in the most general case, it is about 50 μm in order to avoid residual deformation

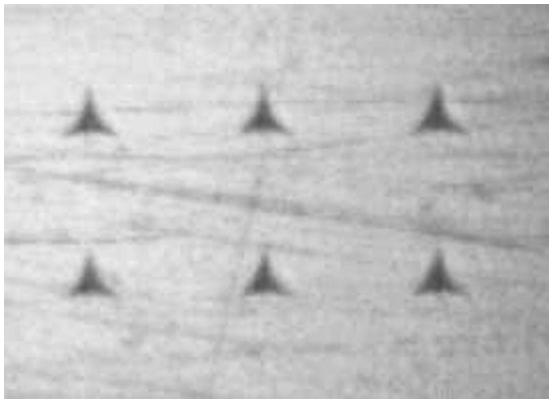


Figure 2. Typical two-dimensional indentation pattern obtained by a Nanoindenter XP™ machine. The distance between the indents is 50 μm.

generated by indentations. This resource permits an easy method to map the mechanical properties in a composite surface, for example. At low loads, the nanoindentation area is small. Then one indentation can be made inside a single grain. In this case, it is also possible to make several indentations, and after that, to make some chemical attack to reveal the grain boundaries. Finally, the indentation position inside or at the grain boundary of grain may be verified, and the different behavior analyzed.

2.2.2. Load x Displacement curve

A typical load versus displacement curve, obtained from the recorded data, is shown in Figure 3. The loading curve presents a typical parabolic behavior which is associated with elastic–plastic deformations during the loading. The maximum tip penetration $h_{\max}$, the contact depth h_c, the residual penetration after unloading h_f, and the measured stiffness S are indicated in Figure 3. The elastic recovery after the unloading corresponds to the difference $h_{\max} - h_f$. The form of these curves, including small changes in the increasing depth rate or in the elastic recovery under unloading, can give information about the surface response to the applied load, and consequently about the mechanical properties at the surface.

2.2.3. Basic Equations

During the indentation process, indenter–sample contact presents elastic and plastic components. Figure 4 shows a schematic view of the indentation contact during loading and after unloading. It was verified by Oliver and Pharr [8] that the unload curve behaves as a power law function

$$P = \alpha(h - h_f)^m \tag{1}$$

where h_f is the final depth after complete unload and P is the load.

Oliver and Pharr [8] showed that the tip contact depth into the sample material h_c at maximum load $P_{\max}$ can be obtained by analyzing the unload curve. Based on the work

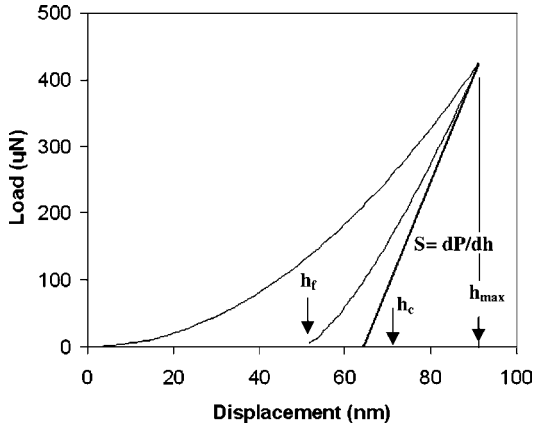


Figure 3. Typical load versus displacement curve showing the loading and unloading parts, the maximum tip penetration ($h_{\max}$), the contact depth (h_c), the final depth h_f, the system stiffness (S), and the elastic recovery.

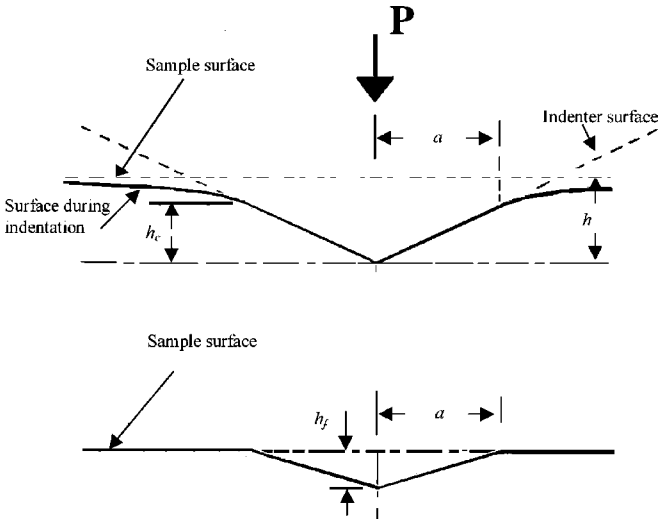


Figure 4. Schematic view of sample surface during and after the indentation, showing the parameters definition used in the equations.

of Sneddon [7] and Doerner and Nix [6], they proposed a method that considers the unload curve as being due only to elastic recovering. The proposed equation, which is the basis for hardness and elastic modulus measurements, is

$$h_c = h_{\max} - \varepsilon \frac{P_{\max}}{S} \tag{2}$$

with S being the unload stiffness,

$$S = \frac{dP}{dh} \tag{3}$$

determined in the initial part of unloading curve. In Eq. (2), ε is a parameter that depends on the tip shape, and is equal to 0.75 for a triangular-based pyramidal Berkovich tip normally used in nanoindentation.

According to the method, the reduced elastic modulus is related to the measured values by the relation

$$E_r = \frac{\sqrt{\pi}}{2} \frac{S}{\sqrt{A}} \tag{4}$$

and

$$\frac{1}{E_r} = \frac{(1 - \nu^2)}{E} + \frac{(1 - \nu_i^2)}{E_i} \tag{5}$$

where the i index indicates the values for the indenter material, and values not indexed correspond to the indented sample material.

The projected contact area is calculated from the contact depth. For a perfect Berkovich indenter, the expression is $A = 24.5h_c^2$ where h_c is the contact depth from Eq. (1). The hardness is then obtained from the usual expression:

$$H = \frac{P}{A} \tag{6}$$

2.2.4. Calibration, Errors, and Limitations

The correct evaluation of mechanical properties from load and depth-sensing indentation depends on the minimization of errors [10] and the comprehension of the method's limitation. However, even in the case of a well-calibrated machine, good results are not obtained if the sample presents a high surface roughness. Another problem appears when the material presents pile up at indentation. Pile up increases the contact area that is not correctly determined when the Oliver and Pharr method is used. Then the calculated values of hardness and elastic modulus are lower than the actual values [11].

The errors and limitations in load and depth-sensing indentation can be separated into three major different groups: calibration of the equipment and determination of contact surface, calibration of the area function for the indenter, and sample effects, such as roughness and pile up. In the following, the calibration processes and how these errors affect the results are detailed.

A correct calibration of the indentation system is very important. Force and displacement calibrations are performed only at installation, and when the system is modified. The following calibration step is to obtain the compliance of the machine. The compliance of an ideal machine for a nanoindentation test must be very low. In this case, the frame machine presents only a small deformation when the load is applied to the sample. However, every machine suffers some deformation during the test. Then, the machine compliance must be deducted from the total compliance obtained from load versus displacement in order to calculate the elastic behavior of the sample. This is done by indenting a sample with high loads, as proposed by Oliver and Pharr [8]. Samples of well-known mechanical properties (like Al and fused silica) are indented at high loads, and the stiffness from the unload is obtained accurately. Considering that the machine is not perfectly rigid, the unload stiffness (dP/dh) has contributions from the elastic responses of both the sample under test and the instrument. The compliance of the instrument $C_f = 1/S_f$ is then incorporated as

$$\frac{1}{S} = \frac{1}{S_s} + \frac{1}{S_f} \tag{7}$$

with S_s being the sample contribution and S_f the machine frame stiffness.

$$\frac{dh}{dP} = \frac{1}{S} = \frac{1}{2E^*}\sqrt{\frac{\pi}{A}} + \frac{1}{S_f} \tag{8}$$

Then, for a Berkovich indenter where $A = 24.5h_c^2$, we have

$$\frac{dh}{dP} = \frac{1}{S} = \sqrt{\frac{\pi}{24.5}}\frac{1}{2E^*}\frac{1}{h_c} + \frac{1}{S_f} \tag{9}$$

Frame compliance is obtained from the data of indentations made at high loads. Values of $1/S$ are plotted versus $1/A^{1/2}$ or $1/h_c$. The data are then extrapolated to an infinite depth or $h_c^{-1} = 0$, and then S_f is calculated. The obtained value is known as the frame compliance of the machine. A common value for load and depth-sensing indentation machine compliance is about 0.1 nm/mN, which corresponds to a frame stiffness of 10^7 N/m.

Environmental effects such as thermal variations and vibrations are another source of errors. Nanoindentation tests have a high sensitivity to temperature changes due to the different values for the thermal expansion coefficients of different materials in the structure frame. Consequently, it is necessary to minimize the thermal fluctuations in the room where the machine is installed. The effect of mechanical vibrations is normally avoided by using a nonvibrating table, and enclosing the machine in an acoustic-damped cabinet.

Thermal drift rates lower than 0.1 nm/s are necessary when small-scale indentations are being performed. Nanoindentation is normally performed at room temperature since the effect of thermal drift is higher when it is necessary to maintain a constant high temperature. However, there are machines suited to operate at temperatures of about 100 °C. In this case, the entire structure composed of the frame, motors, and measurement head are maintained at the same high temperature inside an oven at a constant temperature. This maximum temperature operation is limited by the electrical and electronic components that are inside the oven. Some other high-temperature systems perform nanoindentation at high temperatures, but nanoindentation at high temperatures is still a very complex process.

Another routine calibration is to determine the tip's area as a function of indentation depth. In practical situations, the indenter is not completely sharp, but it shows some roundness at the end, and some errors can be observed because of this fact. Mencik and Swain [10] described a series of errors that are present in the nanoindentation data, which can be due to several factors like tip blunt effects and surface roughness.

The correct determination of the projected contact area as a function of the contact depth is very important to obtain the values for hardness and elastic modulus. Pyramidal indenters present a round end, while spherical indenters are never perfectly spherical. Then, it is necessary to compensate these effects. Scanning electron microscopy and atomic force microscopy performed in real Berkovich indenters showed that indenters with a radius of about 20 nm can be considered relatively good, "sharp" indenters. Berkovich tips extensively utilized have an indenter radius larger than 20 nm. To take into account the blunting of the tip, Oliver and Pharr [8] proposed a relation of the projected contact area to the contact depth that can be described by the equation

$$A(h_c) = 24.5h_c^2 + C_1h_c + C_2h_c^{1/2} + C_3h_c^{1/4} + \cdots + C_8h_c^{1/128} \tag{10}$$

where C_1–C_8 are constant, and determined by fitting the hardness versus penetration data of materials supposed to have a constant hardness at different depths, such as fused silica. Indentations at different loads are performed, and then the area function coefficients are fitted to obtain a constant hardness versus depth profile. This procedure is the most used method to correct the Berkovich rounding tip.

In a recent work, Thurn and Cook [12] presented an analysis of the area function for nanoindentation. A two-parameter area function, obtained from the harmonic average of a spherical tip profile and a perfect conical profile,

was proposed. According to the authors, the two-parameter area function adequately describes the Berkovich indenter tip contact area over the entire load range used in the indentation tests. The calculated projected contact areas for indentations were compared with values obtained from scanning electron microscopy (SEM) observation. Values of the modulus and hardness determined using the two-parameter area function are very similar to the ones obtained using the eight-parameter area function of Oliver and Pharr [8].

The calibrations of spherical indenters are normally made by performing indentations in fused silica in the elastic limit. Then the Hertz elastic model

$$P = \frac{4}{3} E_r R^{1/2} h_e^{3/2} \tag{11}$$

is used to obtain the actual radius of the indenter. However, normally, the radius is not constant, and then some area function is needed to correct the hardness and elastic modulus obtained. A good review of these corrections was presented by Fischer-Cripps [13]. Another calibration process was suggested by Swadener and Pharr [14].

An important factor that limits the determination of the actual surface mechanical properties by nanoindentation is related to the sample surface finishing quality [15–16]. The samples should be flat to have good results under nanoindentation tests. A perfect orthogonal approximation of the tip relative to the surface is more difficult in rough surfaces. In almost all cases of a present roughness, the measured mechanical properties change from point to point in the sample surface, making it difficult to obtain a correct interpretation of the results.

Bobji and Biswas [15–16] analyzed the effect of roughness on hardness values measured by nanoindentation. They established a relation between the penetration depth normalized with respect to a roughness scale parameter. The practical usefulness of their model was verified by the numerical simulation of nanoindentation on a fractal surface. They found that an increase in surface roughness also causes an increased deviation in hardness measured values. They presented a method to deconvolute the effect of roughness in arriving at real hardness characteristics of the near-surface region of a material. The major conclusion they obtained is that, knowing the indenter geometry and given the roughness and penetration depth, it is possible to deconvolute the effect of roughness on measured hardness using a simple algebraic equation, to determine the actual mechanical property profile of the surface region. However, a problem emerges, which is the lack of good knowledge of the indenter geometry.

Other methods to overcome the surface roughness can also be used in order to minimize its influence on the hardness and elastic modulus values: (1) to estimate a correct zero depth surface contact point by analyzing the stiffness of the sample at the contact, and comparing it with similar results in a flat surface [17]; or (2) to determine the hardness based on the stiffness at maximum loading [18], where the contact area is calculated from the loading stiffness at the loading and unloading curves. Using this method and an analysis of the extrapolated contact depth from the loading curve, it is possible to obtain the hardness [19]. In this case, the hardness calculation does not depend on the knowledge of surface contact, and consequently, it may be independent of the knowledge of surface roughness. The usefulness of these two methods, with respect to the direct Oliver and Pharr method, can be observed by comparing the data in Figure 5. The hardness profiles correspond to a Ti nitride surface created by glow discharge at high temperature. A high surface roughness is present. Drastic differences in hardness profiles are observed when the results from the two methods are compared to the one obtained by the Oliver and Pharr method. Both correcting methods agree very well, and the reported values are according to the literature data on titanium nitride hardness.

In a purely elastic contact regime, all materials always sink in during indentation, while for elastic–plastic contact, the material may either sink in or pile up. In some cases, material flowing up from the indentation cavity may occur, as shown in Figure 6. The indenter still maintains contact with the surface, but in this case, the real contact depth becomes larger than the penetration depth obtained from the Oliver and Pharr method. In materials that show pile up, the actual hardness is lower than measured by the Oliver and Pharr method. Bolshakov and Pharr [20] performed a finite-element analysis of the pile-up effect on the measurement of the contact area, hardness, and elastic modulus by load and depth-sensing indentation methods. They found that the parameter $h_f/h_{\max}$, which can be measured experimentally, can be used as an indication of when the pile-up effect is an important factor. According to the authors, pile up is significant only when $h_f/h_{\max} > 0.7$ for materials that do not appreciably work harden. For these materials, if pile up is not considered, the contact area may be underestimated by as much as 60%, and as a consequence, an overestimation of the hardness and elastic modulus occurs. For materials that moderately work harden or in the case of $h_f/h_{\max} < 0.7$, pile up is not a significant factor, and the Oliver and Pharr data analysis procedure can be expected to give reasonable results. The same metal may present different behavior with respect to pile up, depending on the presence of an annealed or cold-worked state. Well-annealed samples present much less pile up than cold-worked samples of the same metal.

Measurements of pile up and sink in of material around Vickers and spherical indentations in metals and ceramics

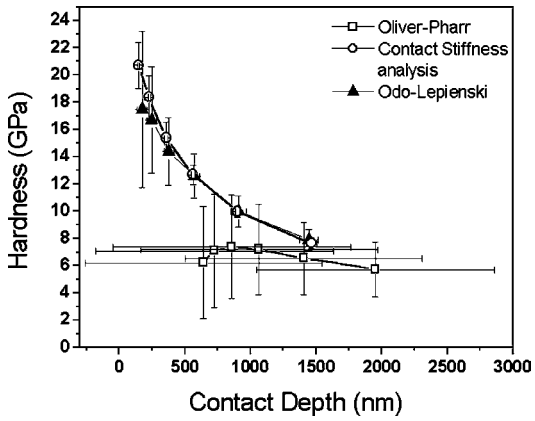


Figure 5. Hardness profile for a high rough Ti-modified surface obtained by using the Oliver and Pharr method [8], contact stiffness analysis [17], and Odo–Lepienski method [19].

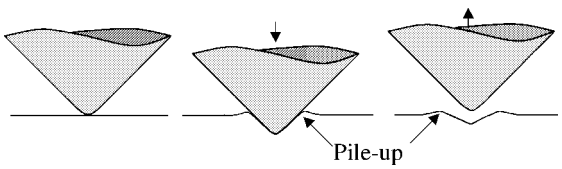


Figure 6. Pile-up phenomena during indentation.

were made by Alcalá and coworkers [21]. They found that surface displacement at the contact boundary under applied load and in the unloaded state is correlated with the uniaxial strain hardening exponent n. Sinking in predominates in materials where $n > 0.2$. A state of nonuniform deformation is detected around Vickers indents, in contrast to the case of spherical indentation, where the deformation state is more uniform as a result of the axisymmetry of the contact conditions. The average (mean) surface deformation state around the contact perimeter in Vickers indents follows a similar correlation with n as that found for spherical indentation. The study of pile up and sink in is very important to the analysis of instrumented indentation experiments.

3. APPLICATIONS

Load and depth-sensing indentation are applied to practically every kind of solid material. This wide field of applications can be attributed to the facility of use since it does not require more than a relatively flat and not rough surface. Comparative results are straight forwardly obtained, being very useful to determine the effect of some parameter variation in the production or treatment of a specimen. In this section, the application and some results of load and depth-sensing indentation in different kinds of materials are presented. The objective is to show the major applications and to discuss the principal results in order to give the reader a direct way to analyze some of the most important fields of nanoindentation application. In addition, some nonconventional and very specific applications are described to show the technique capabilities and new ways to determine the mechanical properties of materials.

Some of the major applications of the nanoindentation technique, as indicated by the number of published articles in the last ten years, are listed: (1) determination of the hardness and elastic modulus of deposited thin films and coatings, principally metallic thin films and diamond-like carbon films; (2) determination of mechanical properties of irradiated materials including ionic irradiated surfaces and plasma-based ion implantation; and (3) study of mechanical properties of brittle materials, such as hardness, elastic modulus, toughness, and adhesion properties.

Some new and not totally well-developed areas where the depth and load-sensing indentation is being applied are: plasticity at low dimensions, mechanical properties of nanostructured materials, stress–strain simulation and residual stress determination, viscoelastic properties of polymers, and the creep and strain rate effect on the deformation of materials at low temperature. In these cases, non-Berkovich indenters may also be used to access mechanical properties other than hardness and elastic modulus.

3.1. Determination of Mechanical Properties of Coatings

In one of the first attempts to obtain the hardness of modified surfaces and coatings using depth and load-sensing indentation, Jonsson and Hogmark [22] developed a useful method for high applied loads based on the tip area function. Some progresses were described by Burnett and Rickerby [23] and Burnett and Page [24]. But independent of all efforts, hardness measurements based on these models are only completely successful for thick layers (on the order of several micrometers) because substrate effects are present, and it is not so clear yet how they can be excluded from the hardness profiles [25–28].

Nowadays, technological advances to produce surface coatings (thin films) and new materials at nanometer scale play an important role in technological applications. The thickness in these situations is reduced to the order of a tenth of a micrometer or less, and consequently, nanoindentation practically is the unique technique indicated to obtain mechanical properties in these regions.

In the following, the principal results of the mechanical characterization of deposited coatings by nanoindentation methods are pointed out.

3.1.1. Variation of Mechanical Properties According to Deposition Process

Coatings are predominantly produced by physical vapor deposition (PVD) and chemical vapor deposition (CVD) processes, which are being continuously developed and improved, with their indication depending on the end product [29–31]. The surface properties are strongly dependent on the bonding nature of the film structure and the substrate [32]. Coatings with an intense covalent bond nature have a tendency to present the highest hardness values, while metals are the lowest, and ionic coatings have intermediate values. Typical applications of coatings are scheduled below simultaneously with their basic composition [33–34]:

- magnetic and electronic films (Nd–Fe–B; Sm–Fe–N; Al–Co; Fe–Co; Cu; Au)
- optical films (SiO_2; TiO_2; Al_2O_3; Ta_2O_5; ITO)
- decorative, hard, and wear-resistant films (Ti–Al–N; Ti–O–N; B–Ti–Nb–N; DLC; Ti–Al–O–N; h–BN; c–BN; Zr–Cu–N; SiC; VC; VN; TiC; WC; AlN; CrAlY)
- corrosion barriers (Nb–Cr; Fe–Cr alloys; Ti–Al–Cr–N–C)
- solid lubricants (MoS_2; C60, V_2O_5, W_3O)

Figure 7 shows the range of hardness and elastic modulus values that can be obtained by using the nanoindentation technique for different material compositions and at different deposition processes as found in the literature. The values present a great dispersion associated not only with the different tailoring processes, but also with different methodologies to obtain the mechanical properties.

Despite the substrate influence on the hardness or film modulus values, typical load/unloading curves can be used to obtain global information about the film-tailoring process. In a comparative study of different coatings, it was recommended to extract these values at a depth on the order of 10% of the coating thickness or, as in some situations, obtain

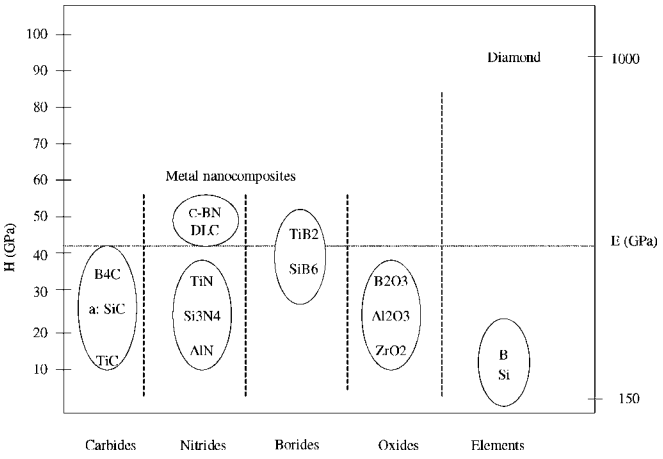


Figure 7. Hardness and elastic modulus of coatings for different deposition processes.

the ones at depths that were reached at the same applied load [25, 28, 36].

Detailed studies of hardness, elastic modulus, and elastic recovery for metal nanocomposite coatings were performed by Musil et al. [36] and Kourtoukova et al. [37]. They showed the relation between hardness to elastic modulus (E/H), hardness to elastic recovery (E^2/H^3), and elastic modulus to elastic recovery (E/W_e and H/W_e) that can be used to tailor the best ensemble of mechanical properties, depending on the end application of the coating.

Literature data are very extensive on the subject of mechanical and tribological properties of coatings using the nanoindentation technique. These properties, referred to before, are strongly dependent on the deposition process used, and the following references are indicated to the reader to be consulted:

- carbides and nitride-based coatings [38–53]
- DLC coatings [54–86]
- metal and oxide coatings [87–92]
- polymeric coatings [93–95]

3.1.2. Effect of Substrate on Coating Characterization

A very large number of articles were written to describe the effect of the substrate in the determination of mechanical properties of thin films [22, 25, 96–105]. Several models were developed to describe the nanoindentation hardness and elastic modulus variation as a function of penetration to film thickness ratio [20, 23, 27, 106–109]. However, this problem continues without a satisfactory solution, and new models are expected to appear in the future. The major difficult factor for modeling the film–substrate behavior under indentation is the complex triaxial stress state in the region constituted by the film, the indenter, and the substrate.

The effect of the substrate in the determination of the hardness of thin films is schematically shown in Figure 8. The elastic zone during indentation is not restricted to the film, but also reaches the substrate. Substrate plastic deformation also occurs for deeper penetrations.

Tsui and Pharr [11] analyzed the effect of hard substrates on the nanoindentation mechanical property measurement

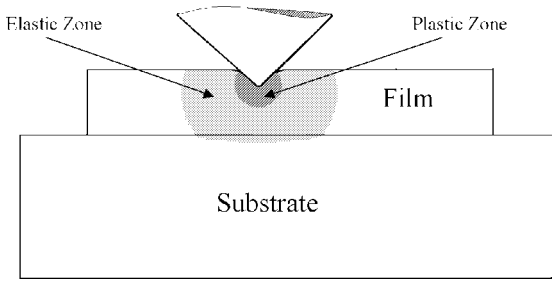


Figure 8. Schematic drawing of the elastic and plastic zone in film substrate.

of soft films. In their work, substrate effects on the measurement of mechanical properties of thin film of aluminum on glass have been studied experimentally by nanoindentation methods. The hardness and elastic modulus of aluminum films with thicknesses of 240, 650, and 1700 nm sputter deposited on glass were systematically characterized as a function of indenter penetration depth using standard nanoindentation methods. They performed scanning electron microscopy and atomic force microscopy of the hardness impressions that revealed that indentation pile up in the aluminum is significantly enhanced by the substrate. It was found the substrate also affects the form of the unloading curve in a manner that has important implications for nanoindentation data analysis procedures. According to the referred article, nanoindentation measurement techniques can overestimate the film hardness and elastic modulus by as much as 100%, depending on the indentation depth, with the largest errors occurring at depths approximately equal to the film thickness. They also verified that indentation pile up in soft aluminum films is significantly enhanced when the films are deposited on hard substrates. In the case where the indentation depth is be about one tenth of the film thickness, the substrate-induced enhancement of pile up is negligible.

Saha and Nix [26] recently performed a study on the effects of the substrate on the hardness and film modulus using the nanoindentation technique. Different films of Al and W were deposited over substrates like Al, Si, glass, and sapphire. The intrinsic hardness and elastic modulus of the films were analyzed using the relation

$$\frac{P}{S^2} = \frac{\pi H}{4\beta^2 E^{*2}} \tag{12}$$

where P is the applied load, S is the contact stiffness, H is the hardness, E^* is the reduced modulus, and β is a tip geometrical constant. According to referred authors, for homogeneous materials, P/S^2 is constant with depth. Then, plotting P/S^2 versus depth may provide useful information about the substrate effect for different combinations of hard film/soft substrate and soft film/hard substrate. The substrate influence is small for soft film over hard substrate. On the other hand, for hard film over soft substrate, the film hardness can be obtained only for total indenter penetrations lower than 10% of the film thickness. Of course, correction of the contact area is needed because pile up will be necessary if the traditional Oliver and Pharr method is used. Considering elastic modulus measurements, the strong effect of

the substrate exists because the elastic field presents a long-range character, and special care must be taken in order to extract the actual elastic modulus value for the film material. For a large mismatch between the film and substrate modulus, King's analysis to estimate the film modulus is indicated [27].

3.2. Surface Modification by Ion Irradiation

Ion implantation (II) and plasma-based ion implantation (PBII) processes are also well established to offer wide possibilities for the surface modification of materials. The application potential of these modified surfaces is very high in the different fields of the modern technology, like the microelectronic, metallurgical, and biological industries. However, these processes are different with respect to coating deposition processes. Atomic species are introduced in the near-surface region by a nonequilibrium process which depends on the parameters used, like ion energy, ion current, chamber vacuum, and temperature conditions. The modified surface region can then be constituted of embedded phases, buried layers, and also, under some conditions, to form a coating layer. In the general case, a graded region is produced from the near-surface to deeper regions. The modified region is, in almost all cases, localized at depths on the order of ten nanometers to thousands of nanometers. Because of this, the nanoindentation technique is the most indicated to obtain mechanical properties such as hardness, elastic modulus, as well as elastic recovery and surface toughness.

Nitrogen is the most used atomic species that has been irradiated to improve surface hardness in metals by II or PBII processes because of its ability to form nitride compounds. However, other atomic species like C, B, Cr, Ti, and O can also be used to prevent wear and corrosion [110]. Notwithstanding, there exist a great number of commercially available metal alloys; the number of reported works in the literature about mechanical properties obtained by the nanoindentation technique for metal surfaces modified by II or PBII is small. Equally as performed for coatings, the influences of the substrate (matrix) and very near-surface region need to be excluded, or at least minimized, in order to obtain the actual hardness and elastic modulus values for the ionic modified surfaces. Surface modifier processes typically produce graded regions. These regions may be formed by a natural oxide layer, the presence of a solid solution of implanted ions, and some stoichiometric phases. The description of a unique hardness value for the material surface is then almost impossible. Consequently, it is recommended to analyze the ionic effect on surface hardening by using the hardness-to-depth profile to compare the different physical parameters of the irradiation process used to modify the surface.

A typical hardness profile for the pure iron surface submitted to N ion implantation is shown in Figure 9. The N peak position is around 70 nm in depth for both substrates. High hardness occurs in the near-surface region, and slowly decreases until it reaches the substrate value. The high surface hardness is attributed to the formation of iron nitrides. The distribution range of the N ions is not sharply defined (it is a Gaussian-like distribution). Then N atoms in solid solution are also present before and after the ion peak

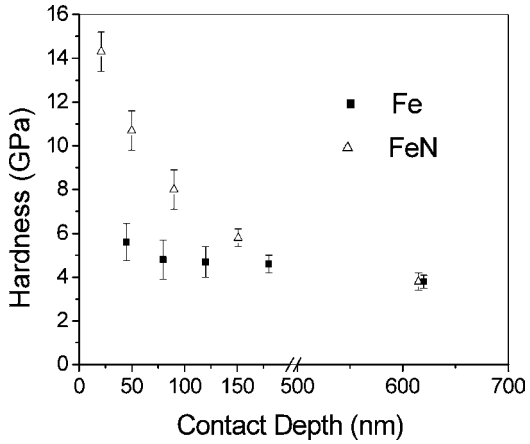


Figure 9. Hardness values as a function of depth for N irradiated iron.

position. Because of the influence of high hardness in this graded region, substrate hardness values are only reached at very deep depths.

Table 1 summarizes the hardness values obtained using the nanoindentation technique (low applied loads $\leq$100 mN), in different commercial metals and metals alloys submitted to different ion irradiation processes (II and PBII).

Additional mechanical property information can also be extracted from the load/unloading curves in the surfaces modified by II or PBII processes. It is well known that II or PBII processes introduce surface damage. The damage profile is basically a function of the ion species and its energy. In this region, atomic displacements, vacancies, Frankel pairs, and extended defects are observed. These defects may alter the elastic field in the region around the tip contact, delaying the plastic deformation by increasing the tip ending pressure, and finally reaching the necessary energy to emit dislocations and form microcracks. This specific phenomenon can be observed in a load/unloading cycle, and it is characterized by a typical tip incursion (pop in) [111]. The extension of the tip incursion at an applied load and time can be then used to estimate the energy to create dislocations. The pattern of the tip incursion allied to electronic microscopy was then used to confirm these hypotheses. At this point, it is important to distinguish this phenomenon from pop in that occurs in dislocation-free materials.

Polymers are also widely used in most industrial applications due to their excellent chemical and physical properties, easy working processes, and low operational costs. On the other hand, when polymers are submitted to load-bearing and abrasive conditions, the mechanical and tribological properties need to be enhanced. The use of II and PBII processes has shown to be a very efficient technique to modify polymer surfaces because there is a large energy transference in II and PBII processes when compared to other more traditional processes like X-ray, γ-ray, ultraviolet rays (UV light), and electron beams. The polymers present a lower bond strength compared to metals, ceramics, or semiconductors; then the transferred energy to the electrons and atoms by the incoming ions stimulates chemical reactions and chain scissions that, at the final stage of the process, produce hardening and stiffening of the polymer surface by a cross-linking mechanism of the chains or

Table 1. Relative hardness (RH) for different metal alloys submitted to different ion implantation processes.

Material	Ion process	Ion type	RH	Ref.
Fe	II	N	3	[111]
		Ar	3	
AISI 1020	II	N	3	
		Ar	3	
		N + Ar	3	
AISI D2	II	N	1.4	[112]
		C	1.8	
AISI M2	II	N	1.48	
		Cr	1.02	
		N + Cr	1.43	
AISI P20	II	C	1.56	
AISI M3:2	II	N	1.51	
		C	1.26	
		Cr + C	1.42	
AISI 420	II	N	1.55	
A286	II	N	1.87	
		Y	1.13	
		Y + N	1.09	
		N + Ar	8.0	
AISI 316L	II	N	1.6	[113]
SS 316L	II	N	2.72	[114]
AISI 304	II	C	1.3	[115]
AISI 440C	II	C	1.15	
		TI	1.15	
AISI 304	PBII	N	2.5	[116]
Al 7075	PBII	O	2.5	[117]
CoCrMo alloy	II	B	1.5	[118]
		N	1.2	
		C	1.3	
		B + N	1.4	
		B + C	1.8	
304L	II	N	1.23	[119]
AISI S7	II	N	1.53	
Ni	II	Ti + C	2.0	[120]
$Ni_{80}Fe_{20}$	II	Ti + C	2.0	
Zr	II	N	4.0	[121]
Al	II	N	5.0	[122]
Al-6061	II	N	5.0	
Ti6Al4V	II	N	3.0–5.0	[123]
Ti6Al4V	PBII	N	10.0	[17]
Ti6Al4V	II	C	2.0	[124]
Si	PBII	C	2.5	[125]
		N	1.1	
Al	II	N	5	[126]

also by the formation of a three-dimensional microstructure which restrains the chain movements.

Literature data about mechanical and tribological properties, using the nanoindentation technique, of polymers, like polyisoquinoline (PIQ), poly-2-vinylpiridine (PVP), polyacrylonitrile (PAN), polyethylene (PE), polycarbonate (PC), polyetherimide (PEI), polystyrene (PS), Kapton polypropylene (PP), polypropylene (PP), polystyrene (PS), polyethersulfone (PES), polyethylene terephthalate (PET), polyhedral oligomeric silsesquioxane (POSS), poly(ether ether ketone) (PEEK), polystyrene (PS), fotoresist AZ 1350, C_{60} films, and plasma-formed polymers, in a bulk or coating form, submitted or not to surface-modifying processes, can be obtained by consulting the references [68, 127–141].

3.3. Hardness at Nanoscale: Anisotropy, Grain Size Effect, and Indentation Size Effect

The study of very small-scale indentation has increased in the last few years. A considerable effort has been made in theoretical and practical aspects of this problem. At very low loads, nanoindentation is being used to understand the strain gradient plasticity and dislocation emission from the plastic region under the indenter.

At larger scales of indentation, a number of models exist for plastic flow, but as experimental scales shrink, properties change; then these models may not be applicable. A detailed understanding of the atomic-level processes that contribute to the initial nucleation of dislocations, their motion, and multiplication will greatly facilitate the design of tailored materials with specific mechanical properties.

Nanoindentation tests performed at very small penetrations increase the probability that some indentations occur only inside one grain in the material, and then anisotropy effects are very important. At high loads, hardness is obtained from a mean value from several grains. At low loads in nanoindentation tests, the hardness value may correspond to one grain only. For polycrystalline materials, the crystalline orientation of the indented grain relative to the tip geometry is not well known. In addition, for well-annealed metals, the typical dislocation separation is about 1 μm, and the region under the tip should behave close to that of a perfect single crystal, that is, a dislocation-free specimen. Consequently, the understanding of the contact mechanics of small volumes becomes increasingly important.

Corcoran et al. [142] developed a study of the anomalous plastic deformation of single-crystal Au(111), Au(110), and Au(100) surfaces under nanoindentation, and observed a yield behavior composed of a series of discrete yielding events separated by elastic deformation. The onset of this behavior is in agreement with calculations for the theoretical shear strength of gold. Good quantitative agreement is found between the experimental results and a model developed for the nucleation and multiplication of dislocations by a simple Frank–Read source under the indenter tip.

In another paper, Fougere et al. [143] used nanoindentation to measure the Young's modulus for nanocrystalline Fe samples produced by inert-gas condensation and warm consolidation. The samples had grain sizes of 4–20 nm and a residual porosity of 2–30% calculated relative to conventional Fe. Values of the Young's modulus for the nanocrystalline Fe are reduced relative to values for conventional, fully dense Fe. According to the above authors, the observed reductions in the Young's modulus for both the nanocrystalline and the conventional porous Fe can be described adequately by several theories utilizing spheroidal porosity.

It is well known that hardness is observed to increase in metals with decreasing indentation size, especially in the submicrometer depth regime [144, 145]. Conventional theories of plasticity do not include material length scales. In these theories, the flow stress at any particular point in a solid is only related to the strain at that point. Some new theories have appeared where the deformation is influenced by the strain gradient present at that point, and related to the concept of geometrically necessary dislocations [146].

In their article, Nix and Gao [146] proposed that the indentation size effect for crystalline materials can be modeled using the concept of geometrically necessary dislocations. Their model describes the variation of hardness with penetration depth as given by

$$\frac{H}{H_0} = \sqrt{1 + \frac{h^*}{h}} \tag{13}$$

where H is the hardness for a given depth of indentation h, H_0 is the hardness in the limit of infinite depth, and h^* is a characteristic length that depends on the shape of the indenter, the shear modulus, and H_0. They observed, from indentation experiments on annealed (111) copper single crystals, cold-worked polycrystalline copper, and single crystals of silver, that this relation is well obeyed. One interesting scale aspect is that indentation hardness can drop by more than a factor of 2 with increasing depth, and this depth dependence appears to be more accentuated in single crystals than polycrystals. Whether this is strictly due to the difference in the initial dislocation density and/or number of sources is still under investigation.

Tymiak et al. [147] studied the plastic strain and strain gradient with nanoindentation at very low penetration depths. Plastic strains and their respective strain gradients produced by nanoindentation have been theoretically interpreted and experimentally measured at shallow indentation depths for spherical and wedge indenters. For a sharp wedge, both experimental continuum-based and theoretical geometrical approaches suggest the strain gradient decreasing with increasing indentation depth. For spherical indentation, theoretical geometrical analysis yields a depth-independent strain gradient proportional to $1/R$. Tungsten and aluminum single crystals exhibit about a factor of 2 decrease in hardness with increasing depth, irrespective of either increasing or decreasing average strain gradients.

Kiely and collaborators [148] performed nanoindentations on Au single-crystal surfaces. They observed two distinct regimes of plastic deformation, which are distinguished by the magnitude of discontinuities in load relaxation. At lower stresses, relaxation occurs in small deviations from elastic behavior, while at the higher stresses, they take the form of large load drops, often resulting in complete relaxation of the applied load. These major events create a relatively wide plastic zone that subsequently deepens more rapidly than it widens. They proposed two regions of plastic relaxation result from two distinctly different processes. In the first region, deformation likely occurs through the nucleation, glide, and locking of several dislocations, resulting in deformation several atomic layers deep. At higher stress levels, dislocation multiplication occurs, and produces mobile dislocations.

3.4. Applications Using Different Indenter Geometries

Several indenter geometries are used in nanoindentation tests. The most common is the Berkovich indenter, which is a three-sided pyramid, with the face angle to normal equal to 65.3°. This indenter has the advantage of being well suited to obtain mechanical properties at a small depth since it is possible to obtain indenters with a relatively sharp end, and at the same time, a good relation contact area to depth. Other indenters commonly used are Vickers, conical, spherical, cube-corner, and the flat-end shape. Each one of these indenters has some advantages and disadvantages. Indenters with different shapes are normally used to obtain mechanical properties diverse from hardness and elastic modulus, although these properties can also be obtained using indenters with a geometry other than Berkovich.

3.4.1. Berkovich

The great majority of results obtained by nanoindentation tests are performed by using Berkovich indenters. A Berkovich indenter has a three-sided pyramidal shape, with angles of 65.3° between the normal and the median of each face, and 76.9° between the normal and each corner line. For a perfectly sharp indenter, the projected contact area is $A = 24.5h^2$, where h is the contact penetration depth. A Vickers four-sided pyramidal indenter has the same area function with respect to the depth penetration relation as a Berkovich indenter. Berkovich indenters are the most used in nanoindentation tests. Although real indenters are never perfectly sharp, three-sided pyramidal indenters such as Berkovich are normally less blunt than four-sided indenters like Vickers.

3.4.2. Vickers

Vickers indenters are also used in load and depth-sensing indentation machines. However, since it is more difficult to obtain sharp Vickers indenters at small depths, its use is more common for some hard materials and higher loads [149]. Some theoretical work also has been performed, comparing Vickers and conical indenters [150, 151].

3.4.3. Spherical Indenters

Spherical indentation was one of the first tests to determine the hardness of material. The Brinnel test is one of the most used to determine the hardness in a macroscale [1]. Francis [152] describes the deformation mechanisms of plastic spherical indentation, in an attempt to correlate the deformation mechanism to the pressure distribution under the indenter as a function of depth penetration.

Spherical indentation has an interesting, and at the same time challenging, aspect: the contact between the indenter and surface is not self-similar, that is, the contact angle increases continuously with penetration. This characteristic permits us to determine properties such as the elastic-to-plastic transition. However, analysis of loading–unloading curves is more complex because the contact angle is varying continuously.

Spherical indenters are widely used, not only to obtain the hardness of the material. A number of papers were published about the application of spherical indentation to determine the hardness and other mechanical properties of materials [153–156].

Depth and load-sensing indentation with spherical indenters is also used to simulate a relationship between the stress and strain, in a similar way as in an uniaxial applied stress, and then to try to obtain the yield strength from appropriate models and nanoindentation load–unload tests. Herbert

and coworkers [157] performed uniaxial tests and nanoindentation tests with spherical indentation in order to explore the accuracy to predict the uniaxial stress–strain behavior of aluminum alloys from nanoindentation tests. They related that spherical indentation can be successfully used to establish an engineering estimate of the elastic modulus and yield strength of Al alloys. However, nanoindentation tests could not reproduce the physical shape of the uniaxial stress–strain curve.

The use of spherical indentations to obtain good results is a more challenging matter. Since spherical indenters are not perfect in the majority of cases, presenting a profile that is not a perfect spherical surface, they are very sensitive to calibration. In most cases, the area and the frame compliance calibrations cannot be obtained from indentations in a unique sample as proposed by Oliver and Pharr for pyramidal indenters. Spherical indenters sometimes show a more acute surface, in others a more blunt surface, and also irregular, not spherical shapes. These effects are not easily incorporated to the area calculus as in the case of the blunt effect of pyramidal indenters. Then each case must be analyzed very carefully. Swadener and Pharr [158] presented a calibration method for nanoindentation with spherical indentation. They proposed to perform indentation on two ceramic materials to the same contact depth. The results are then used to determine the radius of the spherical indenter and the machine compliance. Another analysis of the basic equations and methods to use spherical indenters with load and depth-sensing indentation machines was presented by Fischer-Cripps [13].

Residual stress determination is another application of spherical indentation. Several studies have been performed to determine the residual stress at the surface from indentation techniques. Carlsson and Larsson [159] performed pyramidal indentation in materials with residual stress using pyramidal indenters, and they found that the effect of residual stress on load–unload curves with pyramidal indenters and on measured hardness is small. However, as the contact angle in spherical indentation increases with penetration, the sensitivity to residual stress is higher for spherical indentation than for pyramidal indentation. For these reasons, spherical indentation may be useful to obtain the residual stress at surfaces of metallic materials. A new experimental technique for making measurements of biaxial stress using load and depth-sensing indentation using spherical indenters was presented by Swadener et al. [160]. The residual stress at the surface is obtained from analysis of the mean pressure p_m as a function of the ratio a/R, a being the contact radius and R the indenter radius. The presence of residual stress causes an increases in the values of mean pressure during indentation for the same contact area.

3.4.4. Cube-Corner and Other Sharper Indenters

The indentation of ceramic and other fragile materials can generate cracks, such as radial cracks at the indentation corners and conical cracks. The presence of cracks can be useful to determine the toughness of fragile materials [161, 162]. Then, by measuring the extension of a radial crack at indenter corners, it is possible to determine the toughness of the fragile material. However, radial cracks are not normally observed in indentations with a Berkovich indenter if the applied loads are in the range commonly used in nanoindentation because fragile materials have a critical load for crack initiation. Consequently, cracks are not formed for loads lower than the critical load. It was observed that this critical load is a function of the contact angle between the indenter face with respect to the sample's surface [161]. Increasing the contact angle, the critical load decreases. Loads of about 400 mN are necessary to generate radial cracks at the indentation's corner for Berkovich or Vickers indentation in soda lime glass. On the other hand, a load as low as 20 mN can cause radial cracks when cube-corner indentation is performed in the same material [163, 164]. Because of their low critical load, cube-corner indenters are used in toughness measurements with load and depth-sensing indentation.

Indentation fractures at low loads with a cube-corner indenter are used to study the toughness and adhesion characteristics of thin films [28, 165, 167]. The maximum loads that can be applied by nanoindenters are normally not high enough to induce surface cracks if Berkovich pyramidal indenters are used. However, with sharp indenters, the toughness can be obtained from determination of the radial crack length with different loads and a model proposed by Anstis and Lawn [161, 162]. Some additional difficulties appear when it is necessary to determine the length of cracks because, as loads, the lengths are also very small. The validation of toughness values for thin films cannot be completely established since it can be verified by other methods in the same way as in the case of a bulk material. In thin films, there are also the effects due to the presence of the substrate and the interface region. The characterization of thin-film fragile behavior is a challenging matter since, in nanoindentation tests, it is very difficult to eliminate the effect of the substrate. In addition, it is possible to perform analysis of the adhesion, by qualitatively determining the delamination and interfacial behaviors, with cube-corner indentation. However, the models used to obtain adhesion energy and toughness values need further development.

The use of indenters sharper than the Berkovich indenter has become more frequent recently. The most used of this sharper indenter is the cube-corner geometry, where the angles of the corners are 90°. Cube-corner indenters are normally used in order to evaluate the material's resistance to microcrack initiation and propagation [168] and residual stresses in ceramic materials [169]. In addition, analyses comparing the unload curves obtained with cube-corner, Berkovich, and spherical indentation are also being performed [170]. Nanoindentation results made using indenters with different triangular pyramidal geometries are described by Ikezawa and Maruyama [171].

Malzbender et al. [28], in a review article, described the practical use of nanoindentation and scratch tests in determining the mechanical properties of thin coatings. They presented the major methods to determine the mechanical properties of coatings, with emphasis on brittle coatings.

3.4.5. Flat-End Indenters and Viscoelastic Behavior

The determination of mechanical properties of polymers from nanoindentation tests needs special care due to their viscoelastic behavior. The penetration and load are now

dependent on the strain rate. The indenter penetration depends on the loading time being higher if the test is performed at lower rates. Even at constant load, the penetration depth continues to grow. The values of the hardness and elastic modulus measured by nanoindentation are then dependent on the time and strain rates used in the experiment. For this reason, it does not make sense to define a unique value hardness for a polymer since it depends on the strain rates and interval times used in the experimental procedure. The elastic modulus is also not well defined, the values being obtained from nanoindentation tests higher than those measured from uniaxial tension tests in bulk materials. Much work, however, has been done to determine the mechanical properties in polymers despite this additional difficulty caused by their viscoelastic behavior [172–174].

An additional capability of nanoindentation is to measure the viscoelastic properties of polymers. In order to do so, flat-end cylindrical indenters are normally used. The contact area being constant during penetration, the pressure distribution for a determined depth under the indenter can be considered almost constant if low loads are used. Basic equations of flat punch nanoindentation on polymers are presented by Cheng et al. [175]. The standard three-element viscoelastic material indented by an axisymmetric flat-ended indenter was investigated theoretically. The solutions of the equations of viscoelastic deformation were derived for the standard viscoelastic material for compressible as well as incompressible solids. They analyzed both the flat-punch creep test and the load-relaxation test, providing a fundamental basis for probing the elastic and viscous properties of coatings with nanoindentation tests.

3.5. Scratch Test

A scratch test consists of drawing a surface with a tip applying a normal load (P). Since the nanoindentation system permits very good control in the applied force and penetration depth, it is only necessary to add a tangential movement actuator to obtain a very well-controlled scratch test at nanoscale. The used tip can be conical, spherical, or pyramidal. Theoretical studies have been developed in order to understanding the contact mechanical at nanoscale. Buldan and coworkers [176] presented a very interesting study on the contact at atomic scale of an indenter, and subsequent pulling and dry sliding of a sharp and blunt metal tip on a metal surface. They used molecular dynamics methods and empirical potential on an embedded-atom model to determine the atomic structure evolution and the variation of the normal and lateral forces. The results are very different for each of the diverse tips used in the study. Finite-element analysis of the deformation during indentation and scratch tests on elastic–perfectly plastic materials was performed by Bucaille and Felder [177]. The understanding and quantification of elastic and plastic deformation during a scratch test was presented by Jardret and coworkers [178].

Basically, in a scratch test, the load can be applied in two different ways: (1) a constant load during the entire scratch, or (2) an increasing load (ramp loading) from an initial to a maximum value at the end of the scratch. Lateral force transducers, adapted to a DSI device, are used to measure the lateral force (F) that acts on the scratch tip. If the applied load P tends to zero, and if ploughing is absent, it is possible to estimate the friction coefficient by using

$$\frac{F}{P} = \mu_a + \mu_p \tag{14}$$

where μ_a is the adhesion friction coefficient and μ_p is the ploughing friction [179, 180].

Additional mechanical properties, like elastic modulus, hardness, and fracture toughness, may also be obtained qualitatively by using a scratching test [180–186]. To calculate the hardness, it is necessary to include two components: the scratch hardness H_s and the ploughing hardness H_p [179–185]. For a perfect plastic surface, H_s is obtained by using

$$H_s = \frac{P}{A^*} \tag{15}$$

where

$$A^* = \frac{\pi w^2}{8} \tag{16}$$

is the load-bearing area and w is the local width of the scratch. On the other hand, H_p is defined by

$$H_p = \mu_p \frac{P}{A_p} \tag{17}$$

where A_p is the projected contact area in the normal direction to the tip movement, and is obtained by

$$A_p = R^2 \sin^{-1}\left(\frac{w}{2R}\right) - \frac{w}{2}\left(R^2 - \frac{w^2}{4}\right)^{1/2} \tag{18}$$

with R the tip radius and

$$w = 2\left(2Rh - h^2\right)^{1/2} \tag{19}$$

where h is the tip penetration depth.

Particularly for coatings, the load/displacement curve and its characteristics provide a simple method to obtain the critical load that produces a coating detachment if the scratch device is used in the linear ramp loading mode. The load-displacement curve can show an abrupt discontinuity. Using additional imaging analysis, it is observed that this region corresponds to the onset of the detachment. The critical load to produce cracks is defined by these observations. This process permits us to obtain a measure of the coating/substrate adhesion quality. Of course, the failure modes (cohesive and interfacial) and the load that produces the detachment will depend on the substrate properties, the coating thickness, the tip shape, the loading rate, the friction coefficient, and the surface quality [187–192].

The models to describe the adhesion quality by using the scratch test basically employ the fact that the detachment occurs due to a chipping process on the front side of the tip, with the elastic strain energy realized by an interfacial fracture at the critical load. According to Bull and coworkers [192], detachment occurs when the in-plane compressive

stresses in front of the tip create conditions to induce a critical stress normal to the interface by a Poisson effect. The interfacial fracture energy Γ_i is then obtained by

$$\Gamma_i = \frac{1}{2}\frac{t}{E_c}\left(\frac{\nu_f \mu_c P_c}{A_c}\right)^2 \tag{20}$$

where A_c is the cross-section area at critical load P_c, E_c is the coating elastic modulus, t is the coating thickness, ν_f is the Poisson ratio, and μ_c is the friction coefficient. A modified model of Attar and Johannesson [190] shows that a tangential force is required to remove the coat; on the other hand, Burnett and Rickerby [193] concluded that the elastic–plastic indentation stress is dominant for low friction and thick coats.

Some recent literature reporting scratch tests on the surfaces of different materials, including coatings and bulk, are: for PET films by using a spherical tip at constant load, Beake and Leggett [194]; for fullerene films at constant and ramped load, Lopes [195]; for Cr and CrN films, Hones and coworkers also use constant and ramped loads [196]; a discussion about DLC interface tribology by Menon [197]; Li et al. for metal-particle magnetic tapes [198]; Qi and coworkers for nitrogenated diamond-libe carbon films (CN_x) [199]; Hones and coworkers on hard chromium tungsten nitride coatings [200]; for TiN/SiN_x multilayer coatings, Chen and coworkers [201]; Nelea and coworkers on hydroxyapatite thin films [202]; Charitidis and coworkers, a comparative study of the nanoscratching behavior of amorphous carbon films [203]; investigations of nano- and macrowear of magnetic tape head materials by Tan et al. [204]; in Shen and coworkers, a study of tribological properties of coating/substrate systems in micrometer and nanometer scales [205]; nanotribology studies by Wei and coworkers of Cr, Cr_2N, and CrN thin films using constant and ramped load nanoscratch techniques [206]; evaluation of the adhesion of TiN films using nanoindentation and scratch testing by Toparli and Sasaki [207]; a review of the mechanical properties and tribology of thin sol–gel MTMS coatings by Malzbender and coworkers [28]; indentation depth recovery in a poly(methyl methacrylate) sheet on the microlength scale by Adams and coworkers [208]; nanotribology and surface chemistry of reactively Ti–B–N hard coatings by Ott and coworkers [209]; and hard coating adhesion on ion-implanted polymer using nanoscratch tests by Guzman and coworkers [210].

3.6. Pop-In in Loading Curves

Load-displacement indentation curves in some materials show sudden discontinuities on depth penetration. These discontinuities are called pop ins. A pop-in event may be associated with different aspects, depending on the kind of tested material and its mechanical properties.

A series of articles reporting the applications of nanoindentation to measure the mechanical properties of materials was published, showing the sudden penetration of the indenter, which is followed by elastic or elastic–plastic deformation. Pop ins may occur in very small load indentations made in well-annealed single-crystal materials [211–219]. Other types of pop ins occur during the indentation of hard films deposited over very soft materials [220], and in the case of surface cracking in brittle materials during loading [139].

Pop in also occurs during the indentation of lamellar materials like graphite and mica. Figure 10 shows load versus displacement curves for four different lamellar materials: graphite, mica, GaSe, and NbS_2. Lamellar materials are characterized by highly perfect cleavage, so they readily may be separated into very thin leaves. Indentation tests in these materials present, as a special feature, a deformation process that is practically elastic for pyramidal indenters, even at high loads [221–224]. If the elastic energy is high enough, a sudden plastic deformation occurs and, in most cases, the subsequent deformation may also continue to be elastic.

The mechanical properties of lamellar InSe and GaSe single crystals have been studied by means of nanoindentation tests by Mosca and collaborators [222]. The course of plastic deformation induced in the crystals by application of a definite shear stress through the penetration of a Berkovich tip indicates that the deformation occurs predominantly by pop-in events along easy slip directions, having a fairly elastic character between displacements. Hardness anisotropy along the crystal axes is clearly seen, and the measured elastic modulus presents a discrepancy smaller than 5% in comparison with theoretical calculations performed using previous experimental values of the elastic constants.

Some authors [142] attribute the pop-in excursion to dislocation creation and multiplication by Frank–Read sources. However, in lamellar crystals, which present very different bonding energies at perpendicular directions, the dislocation multiplication by this kind of source does not appear to be the major mechanism to accommodate the plastic deformation during pop-in events. Alternatively, another mechanism that considers the disruption (breaking) of crystal layers and their slip into the existing ones in a layer intercalation process is proposed by Veiga and Lepienski [221].

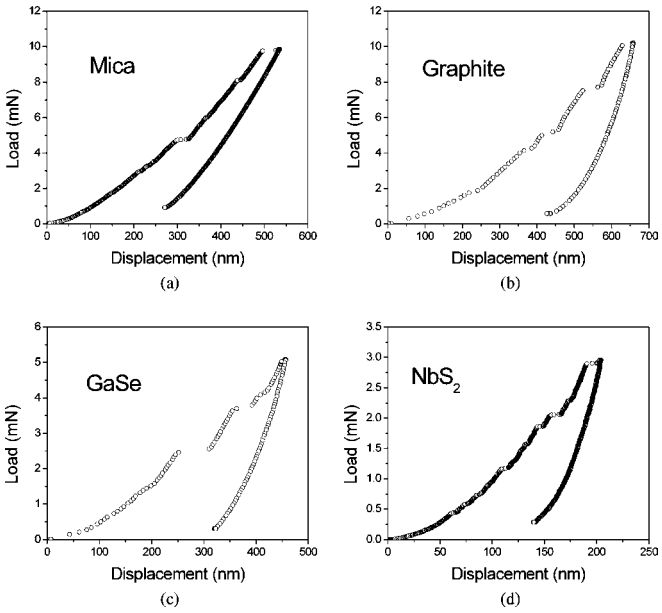


Figure 10. Pop-in events in load versus displacement curves for four different lamellar materials. (a) Mica. (b) Graphite. (c) NbS_2. (d) GaSe.

Gouldstone et al. [211–212] reported nanoindentation experiments followed by TEM and AFM observations on single-crystal and polycrystal Al thin films and polycrystal Cu thin films on Si substrates. They concluded that both single-crystal and polycrystal thin films exhibit periodic displacement pop ins at essentially constant loads during load-controlled nanoindentation. The first displacement pop in appears to occur when the maximum shear stress at the indenter tip is on the order of the theoretical shear strength of the material. The nanoindentation response of the film in between the displacement bursts was found to be purely elastic, and representative of the behavior under a "sharp" Berkovich indenter.

The study of pop-in events can be very helpful to understand the deformation mechanisms that occur under the indenter as the load is applied.

4. THEORETICAL ANALYSIS OF DEPTH-SENSING INDENTATION

Considerable efforts have been undertaken to develop comprehensive theoretical and computational models, in an attempt to elucidate the mechanics and mechanisms of indentation, and to characterize the mechanical properties [225–229]. These models normally are developed based on elastic–plastic models.

Advances in modeling are being obtained by interactive studies, where simulations are verified by experiments and vice versa. The computational simulation of indentation by finite-element methods has been extensively used since the 1970s [230–236].

The first theoretical analysis of indentation was made by Hertz, who examined the elastic contact between two spherical solids [237]. In the context of P–h measurements, for the case of a flat surface (with infinite radius of curvature) which is indented by an elastic sphere of radius R, Hertz showed that

$$P = \frac{4}{3} E_r R^{1/2} h^{3/2} \tag{21}$$

where P is the load, h is the displacement, E_r is the reduced elastic modulus, and R is the indenter radius.

Analyzing the plastic deformation of a spherical indentation, Meyer [1] proposed an empirical relation

$$P = \frac{Ka^m}{(2R)^{m-2}} \tag{22}$$

where m is the hardening factor, a is the contact radius, R is the indenter radius, and K is a material constant. Considering a stress–strain relationship given by

$$\sigma = \sigma_0 \varepsilon^n \tag{23}$$

where σ is the stress, ε is the strain, and n is the strain hardening, Tabor [1] correlated the strain hardening n to the m hardening factor of Meyer by

$$m = n + 2 \tag{24}$$

Johnson [238] suggested that the outcome of a sharp indentation test on elastic–plastic materials will fall into one of three levels, depending on the parameter Γ

$$\Gamma = \frac{E \tan \beta}{(1 - \nu^2)\sigma_y} \tag{25}$$

where E is the elastic modulus, ν is the Poisson coefficient, σ_y is the flow stress at first yield, and β is the angle between the sharp indenter and the nondeformed surface of the material.

According to Johnson, for $\Gamma < 3$, considered level I, very little plastic deformation occurs during the indentation, and the properties can be derived from an elastic analysis. In level II, for $3 < \Gamma < 40$, an increasing effect of plasticity is observed. In this region, both the plastic and elastic properties of the material must be taken into account. Based on the solution of the expansion of a spherical cavity from Hill, and considering the fact that the pressure in the region beneath the indenter is mostly hydrostatic, Johnson suggested that the mean pressure, and consequently the hardness, could be expressed by

$$H = \frac{2}{3}\sigma_y \left(1 + \ln \frac{E \tan \beta}{3(1 - \nu^2)\sigma_y}\right) \tag{26}$$

For $\Gamma > 40$, the plastic deformation is present all over the contact area, and the elasticity no longer influences the hardness value of the material. In this region, the hardness can be given by relation

$$H = C\sigma_y \tag{27}$$

where C is a constant that only depends on the geometry of the sharp indenter. Sometimes σ_y is substituted by σ_r, which corresponds to the flow stress at a plastic strain ε_r. The observed value of C varies from 2.5 to 3 for metals.

Hainsworth et al. [239] observed that the loading curve obtained by depth-sensing indentation experiments can be described by

$$P = Kh^2 \tag{28}$$

where K is a constant. This relationship was tested by finite-element calculation by Zeng and Rowcliffe [240]. The dimensional analysis of Cheng and Cheng [241] supports this assumption. According to Hainsworth et al. [239]

$$K = E\left(\phi\sqrt{\frac{E}{H}} + \psi\sqrt{\frac{H}{E}}\right)^{-2} \tag{29}$$

where E is the elastic modulus, H is the hardness, and ϕ and ψ are empirical constants. The values obtained from Hainsworth were $\phi = 0.194$ and $\psi = 0.903$.

Malzbender et al. [242] derived an analytical expression for the indentation load–depth relation during loading, given by

$$P = E_r\left(\frac{1}{\sqrt{24.5}}\sqrt{\frac{E_r}{H}} + \varepsilon\sqrt{\frac{\pi}{4}}\sqrt{\frac{H}{E_r}}\right)^{-2}(h + \xi)^2 \tag{30}$$

where E_r is the reduced modulus, ε is the geometrical constant which take a value of 0.72 for a conical indenter and 0.75 for a paraboloid, and ξ is a correction for blunt tips.

Larsson [243] developed an investigation of the contact of a sharp indenter under rigid-plastic conditions using theoretical and numerical approaches. He found that, in a general situation, where stress–strain relations are not idealized, there is no single representative value of the uniaxial stress–strain curve that can be used in order to evaluate the global parameters at contact.

There is an important effort underway in the theoretical analysis of nanoindentation. However, analytical models are not available due to the complexity of the problem. Then, finite-element methods are the most promising ways to analyze the contact of a hard indenter with a solid material.

5. CONCLUSION

Nanoindentation is a powerful technique to determine the surface mechanical properties in the nanoscale regime in practically all kinds of solid materials. This technique is especially important for new materials developed by nanoscience and nanotechnology applications. The method is based on the analysis of the load–unload curve, which allows us to obtain information about the elastic–plastic surface behavior in a well-controlled way.

Hardness and elastic modulus values are measured using known analysis methods which are widely accepted nowadays. However, certain additional care is necessary in measurements performed under extreme conditions, such as when the elastic–plastic behavior is strongly dependent on the ratio H/E, or when soft surfaces show pile up around the indenter and very hard surfaces present a sink-in effect. In these situations, the measured values of the hardness and elastic modulus may show deviations from the actual values. Additional care is also recommended for measurements on thin films where the substrate influence is an important factor.

The more common indenters used in nanoindentation are the Berkovich and Vickers tips, but additional information about the mechanical surface properties may be accessed by other kinds of indenters:

- Surface fracture toughness for brittle materials can be studied by using a cube-corner indenter.
- Viscoelastic properties, typical for polymer materials, can be obtained by using cylindrical or flat punch indenters.
- Surface adhesion quality between the coating substrate and residual stress at the surface can be studied by using spherical indenters.

A scratch test performed by using a modified nanoindentation machine can also give extra information about the mechanical surface properties, especially for coatings. However, it is not very clear because of their complexity, and mainly due to the frictional force. The obtained results are qualitatively correct, but more theoretical work is needed in order to compare the different results.

Despite the great number of published papers relating mechanical surface properties acquired by using nanoindentation, supplementary efforts still need to be performed on the theoretical approach in order to quantify and better understand the results in the nanoscale regime.

As a future challenge, it is necessary to develop nanoindentation devices that permit us to perform indentations at higher temperatures. In some materials, temperature dependence on their ductile-brittle behavior occurs, and it may be better investigated if high-temperature nanoindentation is available in the future.

GLOSSARY

Chemical vapor deposition (CVD) process used to product coatings or thin films by using chemical route.

Diamond-like carbon (DLC) a hard noncrystalline carbon film grown by CVD or related techniques, and that contains predominantly sp^2 carbon–carbon bonds.

Friction coefficient In tribology, denotes the dimensionless ratio between the tangential friction force and the normal applied force in a two-body interaction.

Hardness Relative resistance of a metal or other material to denting, scratching, or bending. Denotes, for a solid material, the resistance to penetration by other bodies. The equation used to calculate the hardness is the ratio between the applied load and the residual area produced by the penetration ($H = F/A$). Value depends on the method used for measurement.

Ion implantation Process where atomic species are introduced in the near-surface region by a nonequilibrium process which depends on the parameters used, like ion energy, ion current, chamber vacuum, and temperature conditions. The modified surface region can then be constituted of embedded phases, burried layers, and also, under some conditions, to form a coating layer. In the general case, a graded region is produced from the near-surface to deeper regions.

Nanohardness Denotes, for a solid material, the resistance to penetration by other bodies for depths at nanometerscale.

Nanohardness test Hardness measurement based on analysis of loading–unloading load versus depth curve in nanoscale regime. The indenter used can be Berkovich, Vickers, cube corner, flat end, and cylindrical.

Nanoindentation Process of indenting a solid surface by a pointed hard material at very low load using a load and depth-sensing device to obtain the surface hardness.

Nanoindenter Device constructed to perform nanoindentation and capable of measuring hardness and elastic modulus of material surfaces in the nanoscale regime.

Physical vapor deposition (PVD) process used to produce coatings or thin films by using a physical route.

Plasma-based ion implantation Surface modifier process based on two approaches one for conventional ion implantation, and the other to give thermally enhanced diffusion. The plasma sheath around the target expands dynamically, and permits the target surface to be uniformly treated, forming a modified layer.

ACKNOWLEDGMENTS

The authors thank CNPq-Brazil, Capes-Brazil, and Fundação Araucária-PR-Brazil for the financial support, and Maria Fabíola Vasconcelos Lopes for helping in the final revision.

REFERENCES

1. D. Tabor, "The Hardness of Metals," Clarendon, Oxford, U.K., 1951.
2. J. B. Pethica, R. Hutchings, and W. C. Oliver, *Phil. Mag. A* 48, 593 (1983).
3. J. B. Pethica, in "Ion Implantation into Metals" (V. Ashworth, W. Grant, and R. Procter, Eds.), pp. 147–156. Pergamon, Oxford, U.K., 1982.
4. J. L. Loubet, J. M. Georges, J. M. Marchesini, and G. Meille, *J. Tribology* 106, 43 (1984).
5. D. Newey, M. A. Wilkens, and H. M. Pollock, *J. Phys E: Sci. Instrum.* 15, 119 (1982).
6. M. F. Doerner and W. D. Nix, *J. Mater. Res.* 1, 601 (1986).
7. I. N. Sneddon, *Int. J. Eng. Sci.* 3, 47 (1965).
8. W. C. Oliver and G. M. Pharr, *J. Mater. Res.* 7, 1564 (1992).
9. B. Bhushan, in "Handbook of Micro/Nanotribology" (B. Bhushan, Ed.), Chap. 10. CRC Press LLC, Boca Raton, FL, 1999.
10. J. Mencik and M. V. Swain, *J. Mater. Res.* 10, 1491 (1995).
11. T. Y. Tsui and G. M. Pharr, *J. Mater. Res.* 14, 292 (1999).
12. J. Thurn and R. F. Cook, *J. Mater. Res.* 17, 5 (2002).
13. A. C. Fischer-Cripps, *Vacuum* 58, 569 (2000).
14. J. G. Swadener and G. M. Pharr, *Mater. Res. Soc. Symp. Proc.* 594, 525 (2000).
15. M. S. Bobji, S. K. Biswas, and J. B. Pethica, *Appl. Phys. Lett.* 71, 1059 (1997).
16. M. S. Bobji and S. K. Biswas, *J. Mater. Res.* 13, 3227 (1998).
17. G. B. de Souza, Estudo de Propriedades Mecanicas e Tribologicas de Titanio Nitretado por Plasma, M.Sc. Thesis, Universidade Federal do Parana, Curitiba, 2001.
18. W. C. Oliver, *J. Mater. Res.* 16, 3202 (2001).
19. G. Y Odo, Fundamentos de Nanoindentação e Aplicação em Vidros Submetidos a Migração Iônica, Ph.D. Dissertation (Thesis), Universidade Federal do Paraná, Curitiba, 2001.
20. A. Bolshakov and G. M. Pharr, *J. Mater. Res.* 13, 1049 (1998).
21. J. Alcalá, A. C. Barone, and M. Anglada, *Acta Mater.* 48, 3451 (2000).
22. B. Jonsson and S. Hogmark, *Thin Solid Films* 144, 257 (1984).
23. P. J. Burnett and D. S. Rickerby, *Thin Solid Films* 148, 41 (1987).
24. P. J. Burnett and T. F. Page, *J. Mater. Sci.* 19, 845 (1984).
25. G. M. Pharr, A. Bolshakov, T. Y. Tsui, and J. C. Hay, *Mater. Res. Symp. Proc.* 505, 109 (1998).
26. R. Saha and W. D. Nix, *Acta Mater.* 50, 23 (2002).
27. R. B. King, *Int. J. Solids Struct.* 23, 1657 (1987).
28. J. Malzbender, J. M. J. den Toonder, A. R. Balkenende, and G de With, *Mater. Sci. Eng. R* 36, 47 (2002).
29. D. M. Mattox, "Handbook of Physical Deposition (PVD) Processing." Noyes, NJ, 1998.
30. B. Bhushan, *Diamond Rel. Mater.* 8, 1985 (1999).
31. V. Prabhakaran and F. E. Talke, *Wear* 243, 18 (2000).
32. R. Saha, Z. Xue, Y. Huang, and W. D. Nix, *J. Mech. Phys. Solids* 49, 1997 (2001).
33. "Non-Equilibrium Processing of Materials" (C. Suryanarayana, Ed.), Vol. 2. Pergamon Mater. Series, 1999.
34. S. Hogmark, S. Jacobson, and M. Larsson, *Wear* 246, 20 (2000).
35. R. Saha and W. D. Nix, *Mater. Sci. Eng. A* 319, 898 (2001).
36. J. Musil, K. Kunc, H. Zeman, and H. Polakova, *Surf. Coat. Technol.* 154, 304 (2002).
37. G. L. Kourtoukova, C. Demetry, S. Ramanath, R. M. Andrews, D. S. Jacobs, and R. R. Biederman, *Mater. Sci. Eng. A* 276, 58 (2000).
38. J. M. López, F. J. Gordillo-Vázquez, O. Böhme, and J. M. Albella, *Appl. Surf. Sci.* 173, 290 (2001).
39. D.-G. Kima, T.-Y. Seong, and Y.-J. Baik, *Surf. Coat. Technol.* 153, 79 (2002).
40. J. Esteve, E. Martínez, A. Lousa, F. Montalà, and L. L. Carreras, *Surf. Coat. Technol.* 133–134, 314 (2000).
41. J. M. López, F. J. Gordillo-Vázquez, M. Fernández, J. M. Albell, D. Cáceres, and I. Vergara, *Appl. Surf. Sci.* 172, 110 (2001).
42. H. Ichimura and I. Ando, *Surf. Coat. Technol.* 145, 88 (2000).
43. D. Li, Y. W. Chung, M. S. Wong, and W. D. Sproul, *J. Appl. Phys.* 74, 219 (1993).
44. S. M. Aouadi, D. M. Schultze, S. L. Rohde, K.-C. Wong, and K. A. R. Mitchell, *Surf. Coat. Technol.* 140, 269 (2001).
45. L. Rebouta, C. J. Tavares, R. Aimo, Z. Wang, K. Pischow, E. Alves, T. C. Rojas, and J. A. Odriozola, *Surf. Coat. Technol.* 133–134, 234 (2000).
46. M. Tabbal, P. Mérel, M. Chaker, M. A. El Khakani, E. G. Herbert, B. N. Lucas, and M. E. O'Hern, *Surf. Coat. Technol.* 116–119, 452 (1999).
47. M. Diserensa, J. Patscheider, and F. Lévy, *Surf. Coat. Technol.* 120–121, 158 (1999).
48. K. N. Andersen, E. J. Bienk, K. O. Schweitz, H. Reitz, J. Chevallier, P. Kringhøj, and J. Bøttiger, *Surf. Coat. Technol.* 123, 219 (2000).
49. P. Hones, R. Consiglio, N. Randall, and F. Levy, *Surf. Coat. Technol.* 125, 179 (2000).
50. L. Karlsson, L. Hultman, M. P. Johansson, J.-E. Sundgren, and H. Ljungcrantz, *Surf. Coat. Technol.* 126, 17 (2000).
51. N. Kikuchi, M. Kitagawa, A. Sato, E. Kusano, H. Nanto, and A. Kinbara, *Surf. Coat. Technol.* 126, 131 (2000).
52. A. A. Voevodin and J. S. Zabinski, *Thin Solid Films* 370, 223 (2000).
53. Y. Zhou, R. Asaki, W.-H. Soe, R. Yamamoto, R. Chen, and A. Iwabuchi, *Wear* 236, 159 (1999).
54. M. Berger, L. Karlsson, M. Larsson, and S. Hogmark, *Thin Solid Films* 401, 179 (2001).
55. M. A. Baker, R. Gilmore, C. Lenardi, P. N. Gibson, and W. Gissler, *Vacuum* 53, 113 (1999).
56. B. Feng, D. M. Cao, W. J. Meng, L. E. Rehn, P. M. Baldo, and G. L. Doll, *Thin Solid Films* 398–399, 210 (2001).
57. Y. Panayiotatos, P. Patsalas, C. Charitidis, and S. Logothetidis, *Surf. Coat. Technol.* 151–152, 155 (2002).
58. Y. H. Cheng, B. K. Tay, S. P. Lau, and X. Shi, *Surf. Coat. Technol.* 146–147, 398 (2001).
59. C. E. Bottani, R. Checchetto, A. Miotello, and P. M. Ossi, *Surf. Coat. Technol.* 151–152, 151 (2002).
60. M. P. Johansson, N. Hellgren, T. Berlind, E. Broitman, L. Hultman, and J.-E. Sundgren, *Thin Solid Films* 360, 1 (2000).
61. N. Kikuchi, E. Kusano, T. Tanaka, A. Kinbara, and H. Nanto, *Surf. Coat. Technol.* 149, 76 (2002).
62. B. K. Tay, Y. H. Cheng, X. Z. Ding, S. P. Lau, X. Shi, G. F. You, and D. Sheeja, *Diamond Rel. Mater.* 10, 1082 (2001).
63. Th. Malkow and S. J. Bull, *Surf. Coat. Technol.* 137, 197 (2001).
64. J. Vilcarromero and F. C. Marques, *Thin Solid Films* 398–399, 275 (2001).
65. E. Bertran, F. J. Pino, G. Viera, and J. L. Andújar, *Vacuum* 64, 181 (2002).
66. R. Gago, I. Jiménez, I. García, and J. M. Albella, *Vacuum* 64, 199 (2002).
67. V. Talyzin, L. S. Dubrovinsky, M. Oden, and U. Jansson, *Diamond Rel. Mater.* 10, 2044 (2001).
68. C. E. Foerster, C. M. Lepienski, F. C. Serbena, and F. C. Zawislak, *Thin Solid Films* 340, 201 (1999).
69. Pinyol, E. Bertran, C. Corbella, M. C. Polo, and J. L. Andújar, *Diamond Rel. Mater.* 11, 1000 (2002).
70. K. Yamamoto, K. Wazumi, T. Watanabe, Y. Koga, and S. Iijima, *Diamond Rel. Mater.* 11, 1130 (2002).
71. A. Richter, R. Ries, R. Smith, M. Henke, and B. Wol, *Diamond Rel. Mater.* 9, 1 (2000).
72. Y. Setsuhara, M. Kumagai, M. Suzuki, T. Suzuki, and S. Miyake, *Surf. Coat. Technol.* 116–119, 100 (1999).
73. S. Miyake, Y. Setsuhara, K. Shibata, M. Kumagai, Y. Sakawa, and T. Shoji, *Surf. Coat. Technol.* 116–119, 11 (1999).

74. C. Mitterer, P. H. Mayrhofer, M. Beschliesser, P. Losbichler, P. Warbichler, F. Hofer, P. N. Gibson, W. Gissler, H. Hruby, J. Musil, and J. Vlek, *Surf. Coat. Technol.* 120–121, 405 (1999).
75. E. Liu, X. Shi, H. S. Tan, L. K. Cheah, Z. Sun, B. K. Tay, and J. R. Shi, *Surf. Coat. Technol.* 120–121, 601 (1999).
76. D. F. Wang, K. Kato, and N. Umehara, *Surf. Coat. Technol.* 123, 177 (2000).
77. Y. Kusano, Z. H. Barber, J. E. Evetts, and I. M. Hutchings, *Surf. Coat. Technol.* 124, 104 (2000).
78. C. Quirós, R. Núñez, P. Prieto, I. Vergara, D. Cáceres, L. Soriano, G. G. Fuentes, E. Elizalde, and J. M. Sanz, *Surf. Coat. Technol.* 125, 284 (2000).
79. B. Feng, D. M. Cao, W. J. Meng, L. E. Rehn, P. M. Baldo, and G. L. Doll, *Thin Solid Films* 398–399, 210 (2001).
80. S. J. Bull and S. V. Hainsworth, *Surf. Coat. Technol.* 122, 225 (1999).
81. S. J. Bull and A. M. Korsunsky, *Tribology Int.* 31, 547 (1998).
82. S. Logothetidis and C. Charitidis, *Thin Solid Films* 353, 208 (1999).
83. T. Staedler and K. Schiffmann, *Surface Sci.* 482–485, 1125 (2001).
84. D. Cristopher, R. Smith, and A. Richter, *Nucl. Instr. Meth. B* 180, 117 (2001).
85. F. C. Marques, R. G. Lacerda, G. Y. Odo, and C. M. Lepienski, *Thin Solids Films* 332, 113 (1998).
86. H. Sjöström, S. Stafström, M. Boman, and J.-E. Sundgren, *Phys. Rev. B* 75, 1336 (1995).
87. X. Li and B. Bhushan, *Thin Solid Films* 389–399, 313 (2001).
88. R. Enrique, F. Wu, and P. Bellon, *Surf. Coat. Technol.* 150, 1 (2002).
89. P. Hones, M. Diserens, and F. Lévy, *Surf. Coat. Technol.* 120–121, 277 (1999).
90. E. Lugscheider, S. Bärwulf, and C. Barimani, *Surf. Coat. Technol.* 120–121, 458 (1999).
91. L. Guzman, M. Adami, W. Gissler, S. Klose, and S. De Rossi, *Surf. Coat. Technol.* 125, 218 (2000).
92. N. M. Renevier, V. C. Fox, D. G. Teer, and J. Hampshire, *Surf. Coat. Technol.* 127, 24 (2000).
93. F. Benítez, E. Martínez, M. Galán, J. Serrat, and J. Esteve, *Surf. Coat. Technol.* 125, 383 (2000).
94. E. C. Rangel, N. C. Cruz, and C. M. Lepienski, *Nucl. Instr. Meth. B* 191, 704 (2002).
95. E. C. Rangel, N. C. Cruz, M. A. B. Moraes, and C. M. Lepienski, *Surf. Coat. Technol.* 127, 93 (2000).
96. P. J. Burnett and D. S. Rickerby, *Surf. Eng.* 3, 69 (1987).
97. N. G. Chechenin, J. Bottiger, and J. P. Krog, *Thin Solid Films* 261, 219 (1995).
98. M. F. Doerner, D. S. Gardner, and W. D. Nix, *J. Mater. Res.* 1, 845 (1986).
99. B. D. Fabes, W. C. Oliver, R. A. McKee, and F. J. Walker, *J. Mater. Res.* 7, 3056 (1992).
100. S. V. Hainsworth, T. Bartlett, and T. F. Page, *Thin Solid Films* 236, 214 (1993).
101. J. A. Knapp, *Mater. Res. Symp. Proc.* 438, 617 (1996).
102. T. F. Page and S. V. Hainsworth, *Surf. Coat. Technol.* 61, 201 (1993).
103. G. M. Pharr and W. C. Oliver, *Mater. Res. Bull.* 17, 28 (1992).
104. T. Y. Tsui, J. Vlassak, and W. D. Nix, *J. Mater. Res.* 14, 2204 (1999).
105. T. Y. Tsui, J. Vlassak, and W. D. Nix, *J. Mater. Res.* 14, 2196 (1999).
106. H. Gao, C. Cheng-Hsin, and L. Jin, *Int. J. Solids Struct.* 29, 2471 (1992).
107. H. Y. Yu, S. C. Sanday, and B. B. Rath, *J. Mech. Phys. Solids* 38, 745 (1990).
108. Y.-T. Cheng and C.-M. Cheng, *Appl. Phys. Lett.* 73, 614 (1998).
109. K. C. Tang and R. D. Arnell, *Thin Solid Films* 356, 263 (1999).
110. "Modern Tribology Handbook" (B. Bhushan, Ed.). CRC Press, Boca Raton, FL, 2001.
111. C. E. Foerster, C. M. Lepienski, F. C. Nascimento, and F. C. Serbena, *Nucl. Inst. Meth. B* 175–177, 511 (2001).
112. J. I. Oñate, F. Alonso, and A. García., *Thin Solid Films* 317, 471 (1998).
113. H. Pelletier, D. Müller, P. Mille, A. Cornet, and J. J. Grob, *Surf. Coat Technol.* 151–152, 377 (2002).
114. M. Guemmaz, A. Mosser, J.-J. Grob, and R. Stuck, *Surf. Coat. Technol.* 101, 348 (1998).
115. R. J. Bourcier et al. *Nucl. Inst. Meth. B* 59/60, 905 (1991).
116. X. B. Tian, Y. X. Leng, T. K. Kwok, L. P. Wang, B. Y. Tang, and P. K. Chu, *Surf. Coat. Technol.* 135, 178 (2001).
117. M. Bolduc, D. Popovici, and B. Terreault, *Surf. Coat. Technol.* 138, 125 (2001).
118. B. A. Kehler, N. P. Baker, D. H. Lee, C. J. Maggiore, M. Nastasi, J. R. Tesmera, K. C. Walter, Y. Nakamura, and B. M. Ullrich, *Surf. Coat. Technol.* 114, 19 (1999).
119. K. C. Walter, M. Nastasi, N. P. Baker, C. P. Munson, W. K. Scarborough, J. T. Scheuer, B. P. Wood, J. R. Conrad, K. Sridharan, S. Malik, and R. A. Bruen, *Surf. Coat. Technol.* 103–104, 205 (1998).
120. J. A. Knapp, D. M. Follstaedt, S. M. Myers, J. C. Barbour, T. A. Friedmann, J. W. Ager, III, O. R. Monteiro, and I. G. Brown, *Surf. Coat. Technol.* 103–104, 268 (1998).
121. Y. Miyagawaa, S. Nakao, K. Baba, R. Hatada, M. Ikeyama, and S. Miyagawa, *Surf. Coat. Technol.* 103–104, 323 (1998).
122. T. R. Jervis et al., *Nucl. Inst. Meth. B* 72, 59 (1992).
123. J. P. Blanchard et al., *Nucl. Inst. Meth. B* 82, 63 (1993).
124. A. García, J. L. Viviente, F. Alonso, A. Loinaz, and J. I. Oñate, *Surf. Coat. Technol.* 97, 499 (1997).
125. M. Ueda, C. M. Lepienski, E. C. Rangel, N. C. Cruz, and F. G. Dias, *Surf. Coat Technol.* 156, 190 (2002).
126. S. Lucas, J. Chevallier, and N. G. Chechenin, *Surf. Coat. Technol.* 66, 334 (1994).
127. J. C. Pivin, *Nucl. Inst. Meth. B* 59–60, 1153 (1991).
128. J. C. Pivin, *Thin Solid Films* 263, 185 (1995).
129. G. R. Rao, E. H. Lee, R. Bhattacharya, and A. W. McCormick, *J. Mater. Res.* 10, 190 (1995).
130. M. V. Swain, A. J. Perry, J. R. Treglio, A. Elkind, and J. D. Demaree, *J. Mater. Res.* 12, 1917 (1997).
131. G. R. Rao and E. H. Lee, *J. Mater. Res.* 11, 2661 (1996).
132. G. R. Rao, P. J. Blau, and E. H. Lee, *Wear* 184, 213 (1995).
133. J. C. Pivin and P. Colombo, *J. Mater. Res.* 32, 6175 (1997).
134. J. C. Pivin, P. Colombo, M. Sendova-Vassileva, J. Salomon, G. Sagon, and A. Quaranta, *Nucl. Inst. Meth. B* 141, 652 (1998).
135. J. C. Pivin, G. Brusatin, and G. Zalczer, *Thin Solid Films* 287, 65 (1996).
136. D. H. Gracias and G. A. Somorjai, *Macromolecules* 31, 1269 (1998).
137. C. E. Foerster, F. C. Serbena, C. M. Lepienski, D. L. Baptista, and F. C. Zawislak, *Nucl. Inst. Meth. B* 148, 634 (1999).
138. C. M. Lepienski, I. T. S. Garcia, C. E. Foerster, F. C. Serbena, and F. C. Zawislak, *Nucl. Inst. Meth.* 175–177, 668 (2001).
139. C. E. Foerster, I. T. S. Garcia, F. C. Zawislak, F. C. Serbena, C. M. Lepienski, W. H. Schreiner, and M. Abbate, *Thin Solid Films* 411, 256 (2002).
140. C. Klappenich, L. Pruitt, and K. Komvopoulus, *J. Mater. Res.* 17, 423 (2002).
141. E. Yap, D. G. McCulloch, D. R. McKenzi, M. V. Swain, L. S. Wielenski, and A. Clissod, *J. Appl. Phys.* 83, 3404 (1998).
142. S. G. Corcoran, R. J. Colton, E. T. Lilleodden, and W. W. Gerberich, *Phys. Rev. B* 55, 16057 (1997).
143. G. E. Fougere, L. Riester, M. Ferber, J. R. Weertman, and R. W. Siegel, *Mater. Sci. Eng. A* 204, 1 (1995).
144. N. A. Stelmashenko, M. G. Walls, L. M. Brown, and Y. V. Milman, *Acta Met. et Materialia* 41, 2855 (1993).
145. M. S. de Guzman, G. Neubauer, P. Flinn, and W. D. Nix, *Mater. Res. Symp. Proc.* 308, 613 (1993).
146. W. D. Nix and H. Gao, *J. Mech. Phys. Solids* 46, 411 (1998).
147. N. I. Tymiak, D. E. Kramer, D. F. Bahr, T. J. Wyrobek, and W. W. Gerberich, *Acta Mater.* 49, 1021 (2001).

148. J. D. Kiely, K. F. Jarausch, J. E. Houston, and P. E. Russell, *J. Mater. Res.* 14, 2219 (1999).
149. S. Veprek, A. Niederhofer, K. Moto T. Bolom, H. D. Mannling, P. Nesladek, G. Dollinger, and A. Bergmaier, *Surf. Coat. Technol.* 133, 152 (2000).
150. S. Carlsson and P. L. Larsson, *Acta Mater.* 49, 2179 (2001).
151. S. Carlsson and P. L. Larsson, *Acta Mater.* 49, 2193 (2001).
152. H. A. Francis, *J. Eng. Mater. Technol. Trans.* ASME 272 (1976).
153. J. S. Field and M. V. Swain, *J. Mater. Res.* 8, 297 (1993).
154. M. V. Swain and J. Mencik, *Thin Solid Films* 253, 204 (1994).
155. E. R. Weppelmann, J. S. Field, and M. V. Swain, *J. Mater. Res.* 8, 830 (1993).
156. J. S. Field and M. V. Swain, *J. Mater. Res.* 10, 101 (1995).
157. E. G. Herbert, G. M. Pharr, W. C. Oliver, B. N. Lucas, and J. L. Hay, *Thin Solid Films* 398, 331 (2001).
158. J. G. Swadener and G. M. Pharr, *Mater. Res. Symp. Proc.* 594, 525 (2000).
159. S. Carlsson and P. L. Larsson, *Acta Mater.* 49, 2179 (2001).
160. J. G. Swadener, B. Taljat, and G. M. Pharr, *J. Mater. Res.* 16, 2091 (2001).
161. G. R. Anstis, P. Chantikul, B. R. Lawn, and D. B. Marshall, *J. Am. Ceram. Soc.* 64, 533 (1981).
162. P. Chantikul, G. R. Anstis, B. R. Lawn, and D. B. Marshall, *J. Am. Ceram. Soc.* 64, 539 (1981).
163. G. M. Pharr, D. S. Harding, and W. C. Oliver, in "Mechanical Properties and Deformation Behavior of Materials Having Ultra-Fine Structure," p. 449. Kluwer Academic, 1993.
164. D. S. Harding, G. M. Pharr, and W. C. Oliver, *Mater. Res. Soc. Symp. Proc.* 356, 663 (1995).
165. X. D. Li and B. Bhushan, *Thin Solid Films* 398, 313 (2001).
166. X. D. Li and B. Bhushan, *Thin Solid Films* 315, 214 (1998).
167. T. W. Scharf, H. Deng, and J. A. Barnard, *J. Vac. Sci. Technol. A, Vac. Surf. Films* 15, 963 (1997).
168. D. Galusek and F. L. Riley, *Phil. Mag. A* 82, 2041 (2002).
169. K. O. Kese and D. J. Rowcliffe, "Proceedings of the 6th European Conference on Residual Stresses," 2002, Vol. 404, no. 4, p. 855.
170. N. Iwashita and M. V. Swain, *Phil. Mag. A* 82, 2199 (2002).
171. K. Izezawa and T. Maruyama, *J. Appl. Phys.* 91, 9689 (2002).
172. B. J. Briscoe, L. Fiori, and E. Pelillo, *J. Phys. D: Appl. Phys.* 31, 2395 (1998).
173. A. Strojny, X. Y. Xia, A. Tsou, and W. W. Gerberich, *J. Adhesion Sci. Technol.* 12, 1299 (1998).
174. A. Flores and F. J. B. Calleja, *Phil. Mag. A* 78, 1283 (1998).
175. L. Cheng, X. Xia, W. Yu, L. E. Scriven, and W. W. Gerberich, *J. Polymer Sci. B: Polymer Phys.* 38, 10 (2000).
176. A. Buldan, S. Cirani, and I. P. Batra, *Phys. Rev. B* 57, 2468 (1998).
177. J. L. Bucaille and E. Felder, *Phil. Mag.* 82, 2003 (2002).
178. V. Jardret, H. Zahouani, J. Loubet, and T. G. Mathia, *Wear* 218, 8 (1998).
179. S. Jacobsson, M. Olsson, P. Hedenqvist, and O. Vingsbo, in "American Society of Metals (ASM) Handbook" (P. J. Blau, Ed.), p. 430. Am. Tech. Publ. 18, 1992.
180. G. M. Hamilton, *Proc. Inst. Mech. Eng.* 197 C, 53 (1983).
181. B. R. Lawn, *Proc. R. Soc.* 299, 307 (1967).
182. K. L. Johnson, K. Kendall, and A. D. Roberts, *Proc. R. Soc. Lond. A* 324, 301 (1971).
183. J. Malzbender and G. de With, *Surf. Coat. Technol.* 124, 66 (2000).
184. J. Malzbender and G. de With, *Wear* 236, 355 (2000).
185. J. A. Willians, *Tribology Int.* 29, 675 (1996).
186. B. R. Lawn, S. M. Wiederhorn, and D. E. Roberts, *J. Mater. Sci.* 19, 256 (1984).
187. S. J. Bull, *Surf. Coat. Technol.* 50, 25 (1991).
188. M. H. Blees, G. B. Winkelman, A. R. Balkenende, and J. M. J. den Toonder, *Thin Solid Films* 359, 1 (2000).
189. T. W. Wu, *J. Mater. Res.* 6, 407 (1991).
190. F. Attar and T. Johannesson, *Surf. Coat. Technol.* 78, 87 (1996).
191. M. J. Laugier, *J. Mater. Sci.* 21, 2269 (1986).
192. S. J. Bull, D. S. Rickerby, A. Matthews, A. Leyland, A. R. Pace, and J. Valli, *Surf. Coat. Technol.* 36, 503 (1988).
193. P. J. Burnett and D. S. Rickerby, *Thin Solid Films* 154, 403 (1987).
194. B. D. Beake and G. J. Legget, *Polymer* 43, 319 (2002).
195. J. M. J. Lopes, *Nucl. Inst. Meth. B* 175–177, 673 (2001).
196. P. Hones, R. Consiglio, N. Randall, and F. Levy, *Surf. Coat. Technol.* 125, 179 (2000).
197. A. K. Menon, *Tribology. Int.* 33, 299 (2000).
198. X. Li, B. Bhushan, and M. Inoue, *Wear* 251, 1150 (2001).
199. J. Qi, C. Y. Chan, I. Bello, C. S. Lee, S. T. Lee, J. B. Luo, and S. Z. Wen, *Surf. Coat. Technol.* 145, 38 (2001).
200. P. Hones, R. Consiglio, N. Randall, and F. Levy, *Surf. Coat. Technol.* 125, 179 (2000).
201. Y.-H. Chen, I. A. Polonsky, Y.-W. Chung, and L. M. Keer, *Surf. Coat. Technol.* 154, 152 (2002).
202. V. Nelea, H. Pelletier, D. Müller, N. Broll, P. Mille, C. Ristoscu, and I. N. Mihailescu, *Appl. Surf. Sci.* 186, 483 (2002).
203. C. Charitidis, S. Logothetidis, and M. Gioti, *Surf. Coat. Technol.* 125, 201 (2000).
204. S. Tan, V. Prabhakaran, and F. E. Talke, *Trib. Int*. 33, 673 (2000).
205. W. Shen, B. Jiang, S. M. Gasworth, and H. Mukamal, *Trib. Int.* 34, 135 (2001).
206. G. Wei, T. W. Scharf, J. N. Zhou, F. Huang, M. L. Weaver, and J. A. Barnard, *Surf. Coat. Technol.* 146–147, 357 (2001).
207. M. Toparli and S. Sasaki, *Phil. Mag. A* 82, 2192 (2002).
208. M. J. Adams, D. M. Gorman, S. A. Johnson, and B. J. Briscoe, *Phil. Mag. A* 82, 2121 (2002).
209. R. D. Ott, C. Ruby, F. Huang, M. L. Weaver, and J. A. Barnard, *Thin Solid Films* 377–378, 602 (2000).
210. L. Guzman, A. Miotello, E. Voltolini, and M. Adami, *Thin Solid Films* 377–378, 760 (2000).
211. A. Gouldstone, K. J. Van Vliet, and S. Suresh, *Nature* 411, 656 (2001).
212. A. Gouldstone, H. J. Koh, K. Y. Zeng, A. E. Giannakopoulos, and S. Suresh, *Acta Mater.* 48, 2277 (2000).
213. S. Asif and J. B. Pethica, *Phil. Mag. A* 76, 1105 (1997).
214. S. O. Kucheyev, J. E. Bradby, J. S. Williams, C. Jagadish, M. V. Swain, and G. Li, *Appl. Phys. Lett.* 78, 156 (2001).
215. S. O. Kucheyev, J. E. Bradby, J. S. Williams, C. Jagadish, M. Toth, M. R. Phillips, and M. V. Swain, *Appl. Phys. Lett.* 77, 3373 (2000).
216. M. Oden, M. Ljungcrantz, and L. Hultman, *J. Mater. Res.* 12, 2134 (1997).
217. C. Tromas, J. C. Girard, V. Audurier, and J. Woirgard, *J. Mater. Sci.* 34, 5337 (1999).
218. C. Tromas, J. Colin, C. Coupeau, J. C. Girard, J. Woirgard, and J. Grilhe, *Eur. Phys. J. Appl. Phys.* 8, 123 (1999).
219. N. Mattoso, D. H. Mosca, W. H. Schreiner, C. M. Lepienski, I. Mazzaro, and S. R. Teixeira, *Thin Solid Films* 323, 178 (1998).
220. N. G. Chechenim, J. Bottiger, and J. P. Krog, *Thin Solid Films* 261, 228 (1995).
221. W. Veiga and C. M. Lepienski, *Mater. Sci. Eng. A* 335, 6 (2002).
222. D. H. Mosca, N. Mattoso C. M. Lepienski, W. Veiga, I. Mazzaro V. H. Etgens, and M. Eddrief, *J. Appl. Phys.* 91, 140 (2001).
223. C. M. Lepienski, M. S. Meruvia, W. Veiga, and F. Wypych, *J. Mater. Res.* 15, 2069 (2000).
224. M. Goken, M. Kempf, and W. D. Nix, *Acta Mater.* 49, 903 (2001).
225. K. L. Johnson, "Contact Mechanics." Cambridge University Press, Cambridge, 1985.
226. A. Bolshakov, W. C. Oliver, and G. M. Pharr, *J. Mater. Res.* 11, 760 (1997).
227. A. E. Giannakopoulos and S. Suresh, *Scripta Mater.* 40, 1191 (1999).
228. S. Suresh and A. E. Giannakopoulos, *Acta Mater.* 46, 5755 (1998).

229. P. L. Larsson, A. E. Giannakopoulos, E. Soderlund, D. J. Rowclie, and R. Vestergaard, *Int. J. Solids Struct.* 33, 221 (1996).
230. C. H. Lee and S. Kobayash, *Int. J. Mech. Sci.* 12, 349 (1970).
231. M. Imaoka and I. Yasui, *J. Non-Cryst. Solids.* 22, 315 (1976).
232. A. K. Bhattacharya and W. D. Nix, *Int. J. Solids. Struct.* 24, 881 (1988).
233. T. A. Laursen and J. C. Simo, *J. Mater. Res.* 7, 618 (1992).
234. H. J. Gao, C. H. Chiu, and J. Lee, *Int. J. Solids Struct.* 29, 2471 (1992).
235. B. D. Fabes, W. C. Oliver, R. A. McKee, and F. J. Walker, *J. Mater. Res.* 7, 3056 (1992).
236. A. F. Bower, N. A. Fleck, A. Needleman, and N. Ogbonna, *Proc. R. Soc. Lond. A, Math. Phys. Eng. Sci.* 441, 97 (1993).
237. S. Timoshenko and J. N. Goodier, "Theory of Elasticity," McGraw Hill, New York, 1951.
238. K. L. Johnson, *J. Mech. Phys. Solids* 18, 115 (1970).
239. S. V. Hainsworth, H. W. Chandler, and T. F. Page, *J. Mater. Res.* 11, 1987 (1996).
240. K. Zeng and D. Rowcliffe, *Phil. Mag. A* 74, 1107 (1996).
241. Y. T. Cheng and C. M. Cheng, *Appl. Phys. Lett.* 73, 614 (1998).
242. J. Malzbender, G. de With, and J. den Toonder, *J. Mater. Res.* 15, 1209 (2000).
243. P. L. Larsson, *Int. J. Mech. Sci.* 43, 895 (2001).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanomechanics of Nanoscale Materials

Robert Geer

University at Albany, State University of New York, Albany, New York, USA

CONTENTS

1. INTRODUCTION

1.1. Nanotechnology

The fields of nanotechnology and nanoscience represent an intersection of multiple, sometimes diverse, fields of science with a single commonality: manipulation and functionalization of matter at the nanometer-length scale. As defined by the National Nanotechnology Initiative (NNI), a document crafted by the Interagency Working Group on Nanoscience, Engineering and Technology, nanotechnology "... is concerned with materials and systems whose structures and components exhibit novel and significantly improved physical, chemical, and biological properties, phenomena, and processes due to their nanoscale size" [1]. The essence of this statement revolves around so-called emergent properties or processes that are a sole result of the nanoscale engineering of matter. The NNI goes on to state, "Reducing the dimensions of structures leads to entities, such as carbon nanotubes, quantum wires and dots, thin films, DNA-based structures, and laser emitters, which have unique properties. Such new forms of materials and devices herald a revolutionary age for science and technology, provided we can discover and fully utilize the underlying principles."

The fundamental capability to manipulate the nanoscale organization of matter and control its mechanical, electrical, optical, chemical, and/or biological functionality is not a completely novel idea. For example, R. P. Feynman predated the realization of a nanotechnology solution to integrated circuit fabrication by more than four decades. In the late 1950s, Feynman gave his now famous "Plenty of Room at the Bottom" talk, where he presented the idea of a nanometer scale or molecular scale logic device [2]. Such devices are attractive for a variety of reasons. Molecules comprise structures with characteristic length from subnanometer to hundreds of nanometers. Hence, they bridge the dimensional regime between current lithographic structures and atomic sizes. Moreover, the tremendous flexibility in the design and chemical synthesis of molecular structures provides a broad adaptability with respect to potential device functionality, device implementation, and device operation. Virtually every other field of technology development possesses parallel potential with respect to the device miniaturization and integration conceivable via nanotechnology.

The emergent functionality of nanoscale structures and devices has its origins in the unique aspects of quantum mechanics and molecular self-assembly. The quantum-mechanical or wave nature of matter can be exploited to tailor electronic, optical, and chemical properties of a given material system, while the penchant for matter to spontaneously organize at the molecular-length scale provides novel structural templates with which to manipulate the quantum-mechanical boundary conditions that are so critical to controlling quantum phenomena. This is not to say that quantum-mechanical effects are not present in macroscale systems, or that molecular-level self-assembly guarantees new material functionality. Rather, it is the exploitation of the two in concert that opens vast new areas of technological promise.

Current examples of nanotechnological breakthroughs are continuously emerging. The discovery of molecular-scale graphene tubes, dubbed carbon nanotubes, typifies the potential impacts of nanotechnology [3]. By virtue of the propensity of carbon atoms to form threefold coordinated 2-D atomic sheets and the structural perfection of such sheets when rolled into a tubular geometry, entirely new mechanical, chemical, and electrical functionalities have been demonstrated. Such sheets exhibit mechanical properties superior to virtually all macroscopic materials with respect to the combination of strength, flexibility, and light

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (21–46)

weight. Likewise, the perfection of the carbon atomic configuration in such nanotubes has led to the confirmation of the ballistic transport of electrons along the direction of the tube axis. Interestingly enough, the structural details of the "rolling" of the graphene sheet can be exploited to program metallic or semiconducting electronic properties of the nanotube. There is certainly no analogous macroscopic system which possesses such a rich variety of functionality. Hence, this electrical programmability certainly qualifies as an emergent property. Other examples of emergent functionality associated with the nanoscale control of materials include quantum well and quantum dot phenomena exploited for optoelectronic applications of semiconductors, and the rich fields of macromolecular self-assembly that have led to entirely new paradigms for chemical and biological sensor development.

1.2. Mechanics Versus Nanomechanics

As noted previously, the advances promised by the nanoscale material manipulation of matter "... herald a revolutionary age for science and technology, provided we can discover and fully utilize the underlying principles." The "underlying principles" referred to in this statement are those that govern the modification of the conventional understanding of a system as nanoscale effects or phenomena emerge. Consider the case for nanomechanics. The conventional theories for mechanics and the mechanical properties of materials are perhaps the oldest and best studied of all the classical physical sciences. Although it was recognized early on that the microscopic aspects of materials governed their mechanical properties (e.g., the nature of chemical bonding and the microscopic aspects of surface and interface morphology), the validity of so-called continuum elastic mechanics to correctly describe the mechanical properties of materials (even down to micrometer-length scales) has been self-evident for small deformations. However, as the relevant length scale for mechanical deformation of a material system decreases to that of the interatomic or intermolecular spacing, the theories of continuum mechanics are expected to break down. At such length scales, realistic modeling of the mechanical response of a material or structure must take into account the quantum nature of the material. For example, in crystalline materials, the presence of lattice defects, impurities, and grain boundaries must be included to correctly describe the evolution of internal strains resulting from internal and externally applied stresses. This constitutes classical dislocation theory [4]. (The term "classical" is applied to this theory since it treats the crystal imperfection itself as a mathematical singularity, while assuming that the atomic or molecular displacements in the near vicinity of the imperfection are small and sufficiently well described by continuum elastic mechanics.) Of course, there exist entire classes of inelastic mechanical phenomena that are successfully described by empirical or semiempirical classical continuum theories. Simply enumerating them is beyond the scope of this chapter.

In light of the success of such classical continuum theories, it is natural for one unfamiliar with the field to ask whether or not emergent mechanical phenomena at the nanometer-length scale are to be expected at all. Yes and no are both correct answers to this question. The equations of classical elasticity are based on linear response theory; a stress applied to a solid will produce a deformation (strain) linearly proportional to that stress. This linear relationship is a direct result of the electrostatic potential energy of an atom, ion, or molecule as a function of its relative displacement in virtually all bulk materials. For small atomic displacements from an equilibrium position, this potential energy varies quadratically, leading to a linear restoring force (similar to a spring force). As long as this constraint is satisfied, elastic response may be expected, even if the spatial extent of the object is only a few nanometers [5]. However, by tailoring the atomic or molecular interaction forces within a material, as well as its physical-length scale, dramatic departures from elastic behavior may be observed and exploited. In contrast, it has also been shown that certain nanoscale systems can possess linear response behavior at strains well beyond the limits of conventional elastic materials [6].

Thus, there is no question that nanoscale-manipulated materials may "... exhibit novel and significantly improved ..." mechanical properties or phenomena. The challenges in this field of "nanomechanics" encompass the discovery process by which the underlying principles behind such manipulation are investigated and understood.

1.3. Overview of Present Work

The goal of this work is to review critical areas of research in the fields of nanomechanics of nanoscale materials and structures that represent vital efforts in the development of nanotechnology. This document is not a monograph on the theory of mechanics nor is it a compendium of experimental approaches in the study of mechanics. Rather, it focuses on recent experimental and theoretical investigations of emergent mechanical behavior at the nanoscale. Detailed overviews are presented regarding nanomechanical material systems, as well as the development of new instrumentation that will enable the underlying principles of nanomechanics to be more fully elucidated. To optimize the accessibility of this work to the widest possible audience, broader narrative descriptions of the phenomena under consideration are combined with more formal theoretical and experimental summaries. The necessary mathematical definitions of material mechanical properties or summaries of relevant experimental or computational methodologies will be introduced as required. A background in theoretical or experimental materials science or engineering is not required, although a familiarity with the physical principles of mechanics and elasticity will be helpful.

The work is separated into four major sections. The first section comprises a review of newly developed characterization and metrology instrumentation that is enabling the direct measurement of material mechanical properties on the nanometer-length scale. The successful introduction of direct "nanomechanical imaging" is absolutely essential for the degree of fundamental understanding of the mechanics of nanoscale materials necessary for technological exploitation. The second section reviews perhaps the most important nanoscale system with respect to emergent nanomechanical functionality, nanotubes. The third section

reviews the most recent work regarding nanotube composites, and the fourth section details extremely recent work regarding the mechanics of nanoporous materials for applications in nanoelectronics.

To this point, no significant mention is made considering the field of microelectrical mechanical systems (MEMS) or its more recent incarnation, nanoelectrical mechanical systems (NEMS). Traditionally, MEMS has focused on micromachining 3-D structures utilizing the materials of the integrated-circuit (IC) industry. As MEMS has matured, it has provided entirely new experimental platforms from which to test and evaluate the dependence of material and structural mechanical properties as a function of reduced length scale. However, the vast majority of MEMS experimental and theoretical work will not be discussed as conventional elastic continuum models provide an excellent framework for characterizing relevant mechanical properties of MEMS structures.

2. NANOMECHANICAL IMAGING

Macromechanical characterization is an extremely mature field. The basic theory of mechanics was founded on empirical observations that solid bodies exhibit a deformation of some sort (i.e., shape or volume) under the action of applied forces [4]. Standard continuum mechanical theory describes the change in distance between any two points in a solid under the application of an external stress. Assuming linear response, this treatment results in the conventional definitions of elastic constants (bulk modulus, shear modulus, Poisson's ratio, etc.) and the generalized form of Hooke's law. Mechanical characterization beyond elastic behavior is primarily concerned with the onset of inelastic (plastic) deformations and the investigation of material hardness, fracture toughness, and yield (tensile/compressive) strength. Consequently, the study of mechanics is, in large part, the study of deformation. The study of nanomechanics is no different from the previous discussion of nanomaterials illustrated. Recall that the elastic response of an ideal ordered material is directly determined by the local electrostatic atomic, ionic, or molecular potential energy well, the minimum of which is approximately parabolic. If the local atomic, ionic, or molecular mechanical response is preserved in nanoscale structures, it is reasonable to assume that some analogy to "classical" mechanics will exist, and that analogs to classical mechanical measurements will exist. This assumption is the foundation for the fields of nanoindentation, nanotribology, and nanomechanical imaging. These techniques, generally, utilize some type of point contact probe with a nanometer characteristic length scale. Classical mechanics has, for the most part, been used to interpret the results of such experiments, providing empirical confirmation of the propriety of such models. Acoustic, optical, and ultrasonic microscopy techniques have long been the choice for the micrometer-scale characterization of certain surface and subsurface mechanical properties of solids; however, the inherent resolution limitations in far-field lenses and/or coupling media have limited their application at the nanoscale [7]. Nanotribology is an exciting vibrant field and, although intimately related to nanomechanics, is beyond the scope of this review [8]. Nanoindentation is the culmination of decades of empirical research, and is now on the verge of a truly nanoscale characterization technique, although it is typically implemented as a destructive methodology better suited to planar surfaces than nanoscale structures [9]. Since the rising fields of nanoscience and nanotechnology focus on emergent functionality, it is appropriate to restrict detailed discussions to those methodologies that provide not only nanoscale resolution, but that are appropriate for the characterization of nanoscale structures and devices. Consequently, this section will focus on nanomechanical imaging, defined as the cadre of techniques capable of spatially resolving the variation of mechanical properties of nanomaterials, structures, and devices on the nanometer-length scale. Since the mechanics of these techniques use models developed, in part, for nanoindentation, it will be referenced throughout this discussion. The remainder of this section is organized as follows. First, a brief overview is presented of the simple mechanics of a point-probe contact with a surface. This is followed by a brief description of the experimental apparatus used to control such contacts with nanoscale precision, and to exploit them for the extraction of nanomechanical image data. Third, the development and application of a variety of these techniques will be reviewed with respect to the current literature. Lastly, a review of recent results will be presented to demonstrate the potential of nanomechanical imaging and its coming role in the development of nanotechnology.

If an object is of insufficient size or disposition to measure its deformation upon application of an external stress to determine elastic moduli, it is often necessary to apply a point probe to characterize mechanical properties. Hertz was the first to treat such a situation to determine the deformation of two objects in direct contact under an external load [10–11]. (The original problem treated two hemispherical surfaces. A sphere/plane system was recovered by letting one radius diverge to infinity.) The situation and the necessary parameters are outlined in Figure 1. Assuming only elastic deformation (short-range forces) and no interfacial adhesion between the tip probe and sample surfaces, the relation between the contact radius a of the tip/sample contact and the load F is

$$a^3 = \frac{3FR}{4E_r}, \qquad E_r \equiv \left[E_T^{-1}\left(1-\nu_T^2\right) + E_S^{-1}\left(1-\nu_S^2\right)\right]^{-1} \quad (1)$$

where the S and T subscripts denote the sample and tip, respectively. E_r is defined as the reduced Young's modulus.

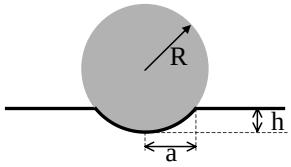


Figure 1. Idealized representation of the elastic deformation of a sphere/plane contact. Note that the radius R of the sphere is defined predeformation.

Taking the origin at the center of the contact area and defining r as the radial distance in the plane of the sample, the elastic indentation h as a function of r is given by

$$h = \frac{3P}{8aE_r}\left(2 - \frac{r^2}{a^2}\right), \qquad r \le a \tag{2}$$

At the apex of the tip, this relation can be rephrased:

$$F = \sqrt{E_r^2 R h^3} \tag{3}$$

The classical derivation of this indentation raises questions regarding its applicability in realistic situations due to the presence of long-range attractive forces (van der Waals), electrostatic forces (surface charging), capillary forces, or general adhesive forces, and not least of all, the quantum-mechanical limit if the mechanical contact is reduced to the atomic-length scale. Several models discussed below include the former effects. And although a number of authors have considered atomic-scale models for tip interaction with respect to the relatively small forces used in topographical imaging, no solely quantum models have addressed mechanical response [12–16]. Although no formal criterion has been universally adopted, it is assumed here (and consistently shown in the vast majority of the experimental literature) that the mechanical response for a surface with a contact radius on the order of 5 nm or larger can be described by elastic mechanics.

If an adhesion energy is associated with the contact area, the local profile of the tip/sample deformation is changed. Johnson, Kendall, Roberts, and Sperling (JKRS) included an adhesion energy w in the Hertzian problem [17–18]. The solid–solid contact deformed locally into a connective neck due to adhesion, resulting in a modified contact area radius:

$$\begin{aligned} \frac{F}{\pi R w} &= a^3 \frac{E_r}{\pi R^2 w} - a^{3/2}\sqrt{\frac{6E_r}{\pi R^2 w}} \\ \frac{hE_r^{2/3}}{(\pi^2 w^2 R)^{1/3}} &= \frac{a^2 E_r^{2/3}}{(\pi w R^2)^{2/3}} - \frac{2}{3}\sqrt{\frac{6aE_r^{1/3}}{(\pi w R^2)^{1/3}}} \\ w &\equiv \gamma_T + \gamma_S - \gamma_{TS} \end{aligned} \tag{4}$$

Here, $\gamma_{T(S)}$ refers to the tip (sample) surface energy, and γ_{TS} is the interfacial energy. Although typically applied to soft materials with large radii, this model has been shown to be applicable over a wide range of materials and nanoprobe tips [19]. An alternate model which reduces artifacts in JKRS theory related to stress variation at the edge of the connective neck was developed by Derjaguin, Muller, and Toporov (DMT), wherein the surface forces are assumed to act only in an annular region just outside the contact radius [20].

$$\begin{aligned} \frac{F}{\pi R w} &= a^3 \frac{E_r}{\pi R^2 w} - 2 \\ \frac{hE_r^{2/3}}{(\pi^2 w^2 R)^{1/3}} &= \frac{a^2 E_r^{2/3}}{(\pi w R^2)^{1/3}} \end{aligned} \tag{5}$$

In all cases, the force–displacement relation depends upon the elastic (or elastic/adhesive) property of the solid. These relationships have been applied to describe the interactions of scanning probe microscope tips and sample surfaces to relate contrast in the scanned image to sample mechanical properties. However, to understand their application, it is necessary to first review the operation of scanning probe microscopes.

The scanning probe microscopy (SPM) platform consists of a nanoscale probe tip placed on or near (∼nm) a surface, and scanned across that surface using a high-resolution piezoceramic crystal scanner (Fig. 2). The probe tip interacts with the surface, and transduces some type of surface-related signal into a digital intensity which is collected into a spatially resolved image with an appropriate raster. The positional control of the piezoscanner is subangstrom, and enables tremendous spatial resolution of the signal transduced by the probe tip. In most cases, the scans are carried out in a feedback mode, where the tip signal is kept constant as the tip is scanned across the sample surface. An early example of such an approach was the so-called scanning tunneling microscope (STM) [21–23]. In an STM, the tip consists of an ultrasharp metal probe which is kept at a reference voltage with respect to a conductive surface. When the tip is brought sufficiently close to the conducting surface (by a piezoscanner), it is possible to measure a current associated with electrons tunneling from the tip to the sample (for separations less than a few nanometers). In this manner, scanning the tip at a constant height permits the measurement of the local density of electronic states (LDOS) at the surface. If the LDOS is constant, the tunneling current between the tip and sample primarily depends on the tip–sample separation, and a feedback circuit can be applied during the scan to keep this current constant by adjusting the tip height during the scan using a vertical piezoceramic scanner. This results in a nanometer-scale image of the surface topography. A separate SPM implementation is the so-called atomic force microscope (AFM) [24] or scanning force microscope (SFM). In the SFM, the nanoprobe typically consists of a pyramidal tip on the end of a micromachined Si cantilever (Fig. 3) [25]. The tip is brought close to the sample. As the tip–sample separation decreases, a long-range attractive van der Waals force deflects the cantilever toward the sample. This is measured via a laser reflected from the backside of the cantilever into a differential photodiode where the deflection of the laser

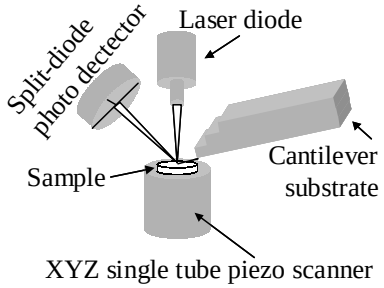


Figure 2. Schematic of conventional SPM configuration. The laser diode beam is focused onto the back of the cantilever at the sample. Deflection of the laser beam due to tip motion is recorded by the split photodetector. This intensity variation is incorporated into a feedback loop incorporating the z axis of the piezoscanner. The x–y axes of the scanner are used to raster the tip across the sample surface.

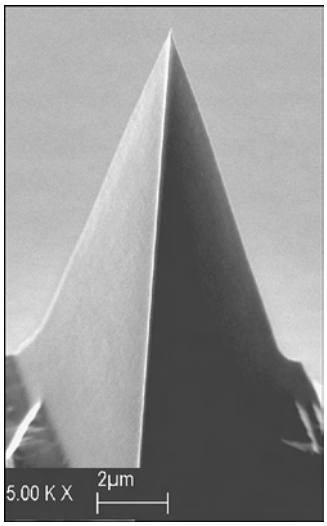


Figure 3. Electron micrograph of pyramidal tip of an Si cantilever.

is calibrated to the change in tip height. If the tip is brought into contact with the sample and scanned across the sample surface, the deflection of the tip can be used to determine the sample topography. Usually, a feedback mode is established wherein the deflection of the cantilever is kept constant by adjusting the relative z position of the tip or sample using a vertical piezoceramic scanner. From this point of view, the end of the pyramidal tip is a nanomechanical probe when a load is placed on the tip in contact with a surface. Although the AFM has been used in a wide range of configurations, we will only focus on those relevant to nanomechanical characterization. An SFM can be used as a nondestructive nanoindentor. However, due to the low spring constant required to provide a high surface topography sensitivity, the tip can only "indent" relatively soft materials. Still, by oscillating the tip below the fundamental resonance of the cantilever against a surface, and measuring the resultant oscillation amplitude as the tip is scanned across that surface, it is possible to record an image for which the intensity is proportional to the mechanical response. This mode is conventionally referred to as force-modulation microscopy (FMM) since the tip–sample force is modulated as the sample is scanned [26–30]. A simple theoretical treatment of this mode was treated by Burnham and co-workers to investigate not only the low-frequency (FMM) mechanical response, but also the high (superresonant) frequency response supplied to the sample by a piezooscillator fixed to the sample [31].

That model is detailed in Figure 4. The sample elastic properties are modeled as a simple spring of spring constant k_i with damping β_i. Assuming that $E_T \gg E_S$, the quantity k_i is linearly related to the sample Young's modulus, and is referred to as the contact stiffness. The details of this relationship are dependent upon the choice of model used. The time dependences of the tip and sample along with the equation of motion are [31–32]

$$\begin{aligned} z &= z_o + z_1\cos(\omega t); \quad d = d_o + d_1\cos(\omega t - \delta) \\ m^*\ddot{d} + 2m^*\beta_c\dot{d} + k_c d &= 2m^*\beta_c(\dot{z} - \dot{d}) + k_c(z - d) \end{aligned} \tag{6}$$

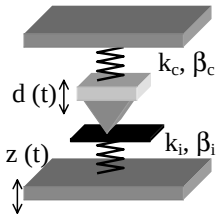


Figure 4. Schematic of tip–sample model used by Burnham et al. The time-dependent vertical position of the sample is denoted by $z(t)$. The resulting deflection of the tip is denoted by $d(t)$. $k_{c,i}$ and $\beta_{c,i}$ denote the spring stiffness and damping constants of the tip and sample, respectively.

The effective mass of the oscillator is defined in terms of the resonant frequency of the cantilever: $\omega_o = (k_c/m^*)^{1/2}$. This is the classic damped-coupled oscillator problem. Solving the equation of motion yields the frequency dependence of the relative displacement ratio d_1/z_1 and the phase δ:

$$\begin{aligned} \frac{d_1}{z_1} &= \frac{k_i\sqrt{1 + (2m^*\omega\beta_i/k_i)^2}}{\sqrt{(k_i + k_c - m^*\omega^2)^2 + 4m^{*2}\omega^2(\beta_i + \beta_c)^2}} \\ \tan\delta &= \frac{2m^*\omega[\beta_i(m^*\omega^2 - k_c) + \beta_c k_i]}{k_i(k_i + k_c - m^*\omega^2) + 4m^{*2}\omega^2\beta_i(\beta_i + \beta_c)} \end{aligned} \tag{7}$$

These results yield two important limiting cases with respect to SFM mechanical imaging modes. Consider the low-frequency ($\omega \ll \omega_o$) and high-frequency ($\omega \gg \omega_o$) limits of this motion:

$$\begin{aligned} \frac{d_1}{z_1} &\approx \frac{k_i}{k_i + k_c}, \qquad \text{for } \omega \gg \omega_o \\ \frac{d_1}{z_1} &\approx \frac{k_i\omega_o^2}{k_c\omega^2}, \qquad \text{for } \omega \gg \omega_o \end{aligned} \tag{8}$$

The low-frequency limit is indicative of the conventional FMM mode. Here, the ratio of the tip deflection (experimentally measured quantity) to the sample oscillation amplitude is only sensitive to the surface mechanical properties if the contact stiffness is on the order of the cantilever spring constant. Hence, if the cantilever is "soft" enough to measure nanoscale topography, it will only be capable of qualifying the spatial dependence of the surface modulus for soft materials. Consequently, FMM is primarily a tool used for polymeric or soft materials. The high-frequency limit is more intriguing from the viewpoint of nanomechanics. At high frequencies, the cantilever cannot follow the driving frequency (superresonant motion), and the driven tip oscillation amplitude decreases as a square of the driving frequency. However, the prefactor of this amplitude depends linearly on the ratio of the surface contact stiffness to the cantilever spring constant. In other words, large deflection implies large, local contact stiffness. Imaging in this mode can provide a signal directly proportional to the surface Young's modulus. Since the contact area of the tip is still on the order of the radius of curvature of the pyramidal tip (~5–20 nm), this mode (christened SLAM by Burnham for scanning local acceleration microscopy) is truly capable of imaging the mechanical properties of a wide range of

materials. Care is required in this technique as the superresonant frequency must fall within the bandwidth of the photodiode detector of the SFM. Also, the expression [Eq. (7)] is valid only in the limit of amplitude-independent contact stiffness, that is, a linear force–displacement response. This is not strictly true in the case of the aforementioned JKRS or DMT models due to the force-dependent area of the contact (even in the case of a flat-end punch) [17, 18, 20]. Consequently, the calibration of SLAM images requires detailed knowledge of the true contact stiffness so that a region of model validity can be determined. However, the simplicity of the model is a major advantage in data interpretation and image artifact identification.

A variation of this approach was investigated several years earlier by Yamanaka et al. [33–34]. Similar to SLAM, an out-of-plane surface vibration was applied to the sample. This approach considered ultrasonic frequencies beyond the response of the photodiode, so the direct oscillation of the cantilever was not observed. However, at high oscillation amplitudes (typically >0.4 nm), a nonzero dc tip deflection was observed. This observation implied that the nonlinear force–displacement curve was being probed, as expected from the models presented above. Moreover, the use of such frequencies reduced the tip oscillation amplitude dramatically, rendering it effectively immobile. As such, the tip constituted (on the high-frequency time scale) an infinitely stiff surface against which the sample periodically is indented [34]. Ideally, then, the force–displacement curve for the tip–sample indentation should precisely correspond to the models presented above. The appearance of a nonzero dc displacement of the tip implies that a nonlinear region of the force–displacement curve is being probed. The oscillation amplitude at which this nonlinearity appears is then related to the slope of the linear portion of the F versus h curve, that is, the contact stiffness. This imaging mode has been referred to as ultrasonic force microscopy or UFM.

The critical concept for mechanical imaging via UFM is the inherently nonlinear interaction of a scanning-probe cantilevered tip in contact with a surface that is undergoing out-of-plane ultrasonic vibrations at a frequency far exceeding the resonant frequency of the cantilever [33]. In such a case, the cantilever is inertially damped and, on the ultrasonic time scale, effectively rigid [34]. Hence, the surface rigidity of materials with contact stiffness orders of magnitude higher than the cantilever spring constant can be quantitatively measured. To illustrate how surface nanomechanical rigidity can be extracted via UFM, Figure 5 shows a schematic of a prototypical force–displacement curve for a rigid nanoprobe tip in contact with a surface. Negative displacement (indentation) of the sample by the tip results in a strong repulsive force. For positive displacement, an attractive force exists between the tip and sample until the sample is pulled away from the tip. This attractive force may result from van der Waals forces, adsorbed fluid layers, or other types of adhesive interactions.

For a scanning probe tip in static contact with a surface, the relative tip–sample distance is h_o and the tip–sample contact force is F_o. Apply an ultrasonic out-of-plane oscillation $\Delta h(t) = h_o + A\cos(\omega t)$. If A is small, the average of the force excursion ΔF is approximately equal to F_o, that is,

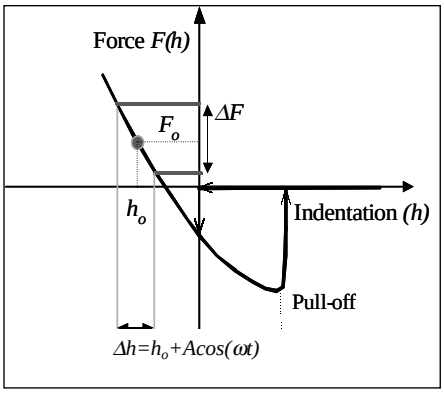


Figure 5. Schematic force–displacement curve for nanoprobe tip and a sample surface. $F(h)$ is the applied force. h is the relative displacement between the tip and the undeflected sample surface.

the tip undergoes no average displacement. If the oscillation amplitude is increased, the nonlinear region of the curve must be included to correctly calculate the average force on the cantilever tip [35]:

$$\langle F(h_o, A)\rangle = \frac{\omega}{2\pi}\int_0^{2\pi/\omega} F(h_o - A\cos(\omega t))\,dt \qquad (9)$$

$$k_c z_c = \langle F(h_o, A)\rangle \qquad (10)$$

where k_c and z_c are the cantilever spring constant and tip deflection, respectively. A simulation of this ΔF as a function of oscillation amplitude A is shown in Figure 6 using the JKRS model [17–18]. The inset in Figure 6 highlights a critical feature of this average force. There is a "kink" or slope discontinuity in the averaged force, which appears for oscillation amplitudes exceeding a certain threshold value A_{th}, for which the tip approaches pull off. This slope discontinuity in $\langle F(h_o, A)\rangle$ translates into a sudden increase in steady-state tip deflection [19]. The threshold amplitude is directly related to the slope of the force–response curve, defined as the contact stiffness between the sample surface and tip

$$S(h) \equiv \frac{\partial F}{\partial h} \qquad (11)$$

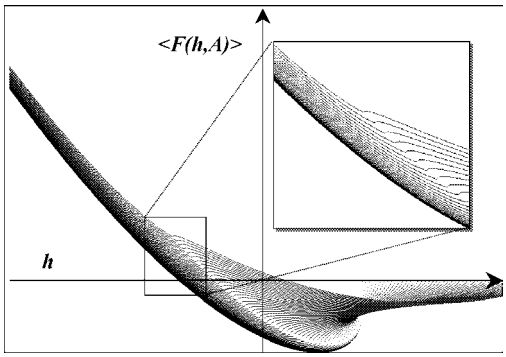


Figure 6. Simulated average tip–sample force–displacement curves for increasing values of A using Eq. (9) and the JKRS model (normalized units). Inset details the discontinuous slope resulting from oscillation amplitudes approaching or exceeding tip pull off.

In the repulsive regime of the force–displacement curve, S is related to the reduced Young's modulus of the tip and the sample. The details of this relationship depend on the complexity of the model used to describe the tip–surface interaction. For a simple Hertzian contact [10–11]

$$S_{\text{Hertz}} = \sqrt[3]{6\overline{F}RE_r^2} \tag{12}$$

$\overline{F}$ is the average force, and R is the scanning probe tip radius. For a more realistic approach, such as the JKRS model, the relationship is more complex, and can be extracted using a forward modeling approach based on curves similar to those in Figure 6.

As noted above, it is critical for the oscillation frequency to *significantly exceed* the fundamental cantilever resonance frequency (typically hundreds of kilohertz). In this frequency domain, the cantilever is inertially damped, not responding quickly enough to the oscillation to be deflected. For cantilever probes designed for use with conventional SPMs, this superresonant frequency typically exceeds 1 MHz [33]. The indentation of such tips at the forces of interest are totally elastic, and do not damage the sample. Frequencies as high as 80 MHz have been used, although higher order cantilever flexural resonance effects must be considered or instabilities may arise [36–37].

Since A_{th} depends monotonically upon the sample contact stiffness, its spatial variation will provide a map of the surface elastic response. Periodic ramping of the vibration amplitude beyond A_{th} as the tip is scanned across the sample provides elastic maps in the same fashion as topographic images. Figure 7 shows a simulated average tip deflection resulting from the application of a triangular-modulated vibration amplitude to the tip–sample displacement [36]. In the half cycle containing the ramp, note the discontinuity at A_{th}. Experimentally, this periodic tip deflection is extracted for modulation frequencies that are large compared to the tip–displacement feedback frequency in a normal topography mode, but small compared to the response frequency of the SPM photodiode detector used to measure tip deflection (typically, a few kilohertz). Experimental results are shown in Figure 8 for UFM tip deflection responses from aluminum and polymer substrates [38]. The differentiation of threshold voltages between the two materials is clearly

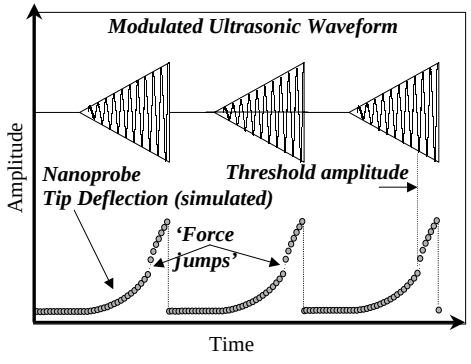


Figure 7. Simulated deflection response of a cantilevered tip (bottom) to a modulated ultrasonic vibration (top). Note that the onset of tip deflection requires a threshold amplitude.

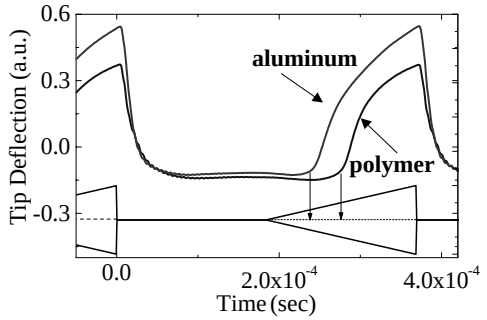


Figure 8. Experimental ultrasonic tip-deflection curves obtained from aluminum (upper curve) and polymer (lower curve) surfaces. Arrows denote the onset of tip deflection, estimating threshold amplitudes. The ramped modulation envelope for the ultrasonic sample vibration is shown at the bottom of the figure.

resolved. To translate these tip-deflection curves into pixilated mechanical image maps, lock-in amplification is used to translate the area under each curve into a raster-scan pixel intensity.

Determination of the sample modulus from these data can be carried out by mapping the amplitude response of the tip deflection to extract the functional dependence of the threshold amplitude versus F_o [33, 35, 39]. Extracting the amplitude variation of the "kink" (Fig. 7) provides the necessary information to determine the static or zero-amplitude contact stiffness from which the Young's modulus can be calculated. Figure 9 displays the ultrasonic tip-deflection response as a function of applied tip force F_o for a spin-on polymer film. The inset in Figure 9 plots the variation of threshold amplitude against increasing force. Assuming a Hertzian response for the contact stiffness, the reduced Young's modulus was directly calculated to be 4.8 GPa using a known tip radius and the absolute vibration amplitude (determined from an optical vibrometer). Using this methodology, Young's modulus measurements have been undertaken for materials ranging from nanoporous silicates ($E_r \sim 2$ GPa) to Si_3N_4 ($E_r \sim 300$ GPa). Application of this technique by Dinelli et al. reported relatively large sample-to-sample variations of modulus measured using this method [35]. This variation is significantly reduced for measurements carried out in vacuum following a low-temperature bake to remove surface-adsorbed water

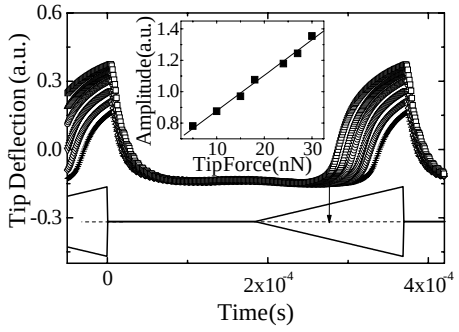


Figure 9. Variation of ultrasonic tip deflection as a function of applied tip force F_o for a polymeric film. The variation of threshold amplitude is apparent. Inset plots this variation as a function of applied force.

films. For silicon oxide films, recent work has demonstrated an absolute error for modulus determination of <25%. Relative local variations on a single sample (point to point) are less than 1% [40].

Burnham et al. described this mode as the "mechanical diode" mode in that the deflection response to an imposed vibration was not symmetric [31]. UFM has been used by that group and a host of others to characterize multiple nanoscale systems to investigate the nanomechanical variations of materials. One of the earliest applications of UFM was by Yamanaka et al. to image subsurface edge dislocation lines in highly oriented pyrolytic graphite [34, 41]. Their work revealed near-surface edge dislocations in the UFM scan as low-contrast lines against a bright field. No associated features were evident in the topography. The subsurface sensitivity of UFM (as for FMM) is attributed to the elastic deformation region in the vicinity of the tip/sample contact. Using a simple static picture, it is reasonable to assume that defects in the near-surface region would alter the local effective modulus. This is the picture commonly used for nanoindentation (usually for a much larger tip/surface contact area). A more accurate picture for the UFM case is one of wave propagation. In the reference frame of the sample, the tip impinges with ultrasonic frequency, and the average displacement of the tip depends upon the surface boundary conditions and the transfer of acoustic energy [42]. A localized defect below a solid surface will affect the average tip displacement if it is within the volume probed by the evanescent acoustic wave excited in the sample. Initial estimates placed these defects within approximately 10 nm of the surface. More recent work by the same authors have not only used UFM as a probe for such subsurface edge dislocation lines, but have also used the pressure field below the tip to induce dislocation motion. By observing the dislocation as a function of load, they were able to displace a single dislocation by 20 nm for a load increase of 80 nN. Reduction of this load permitted the defect to move to its original location. The observation of such dislocation motion is a truly remarkable achievement in nanomechanics. Burnham et al. also used UFM to image subsurface mechanical variations. In that case, a thin polyamide film was applied over a GaAs grating (250 nm period). The polyamide thickness was not revealed in [31] so the depth sensitivity could not be inferred.

Geisler et al. specifically considered sensitivity of UFM to large subsurface voids in Si [43]. The experiment consisted of ion milling angled voids in the cross section of a single-crystal Si wafer. A void ceiling height varied linearly, providing a controlled void depth. From the point of view of mechanics, the deflection of the Si as the tip was scanned across the region resulted from deflection rather than deformation (compression). The contact stiffness of the surface was modeled using finite-element analysis (FEA) to obtain an effective deflection map. Initial investigations assigned a maximum detectable microvoid depth of ~200 nm. However, the large contact forces required in this method often resulted in significant wear of the tip [43].

Altemus and co-workers similarly utilized the UFM to map the deflections of suspended microstructures (cantilevered mirrors) to investigate oscillation-induced hardening of the Au–Cr–Si trilayer beam [44]. After prolonged oscillation ($\sim 10^9$ cycles), slight gradations in the UFM intensity were observed that presumably resulted from high-amplitude induced wear. The gradations most likely reflect variations in the effective high-frequency contact stiffness as subresonant effects would most likely not be observable via UFM.

The capability of UFM to investigate variations of the surface elasticity of higher modulus materials was demonstrated for Ge quantum dots grown (QDs) grown on a silicon surface. Kolosov and co-workers carried out a UFM investigation of such a system [45]. QDs were deposited following Kaimins [46]. Their data showed a clear UFM contrast difference between the Si ($E_{\text{bulk}} = 164$ GPa) and Ge ($E_{\text{bulk}} = 121$ GPa). More recently acquired data on the same structures are shown in Figure 10 [47]. As in [45], the quantum dots possess a relatively unimodal diameter distribution peaked near 100 nm. The QD height was typically 15 nm (~50 Ge monolayers). Closer inspection of the UFM image in the region between the large Ge QDs reveals lower gradations of elastic contrast. When compared to the topographic image, these areas represent plateaus approximately 0.4 nm in height. We associate these regions with a single or double Ge monolayer [48–49]. The ability of UFM to apparently distinguish the mechanical response of only a few monolayers of Ge is impressive, and an important demonstration of its utility for nanomechanical imaging. Since no comparable studies have been published regarding SLAM on similar structures, it is difficult to gauge the comparative sensitivity of the two.

The application of UFM for other nanoparticle films has recently been examined. An investigation of the nucleation of thin Cu films was carried out to investigate the ability of UFM to distinguish between nanoscale Cu clusters and an underlying TaN film. The Young's modulus for Cu varies between 110 and 120 GPa, depending upon processing [50–51]. E for crystalline Ta is approximately 189 GPa, while for TaN, E can exceed 300 GPa, although typical vapor-phase processed TaN can exhibit mechanical moduli variations of 40% below that maximum [52]. Consequently, it would be expected that UFM imaging of Cu nanoparticles on a TaN substrate would appear as low-contrast islands on a bright background. Figure 11 shows a series of images for

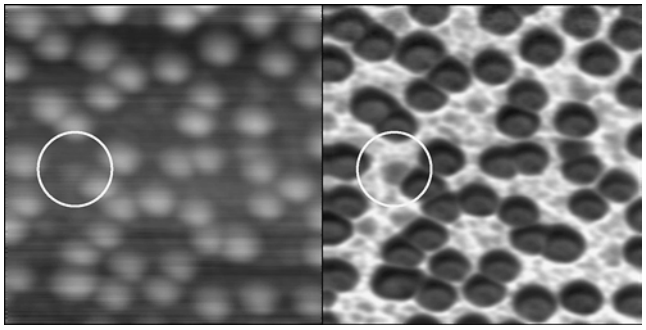


Figure 10. 1 μm × 1 μm AFM scans (left) and UFM scans (right) of Ge QDs on Si. The bright background of the UFM scan corresponds to the higher modulus Si substrate. The gray contrast gradations between the dark 100 nm diameter Ge QDs are tentatively identified as monolayer islands of Ge. One such area is highlighted in both scans by the white circle.

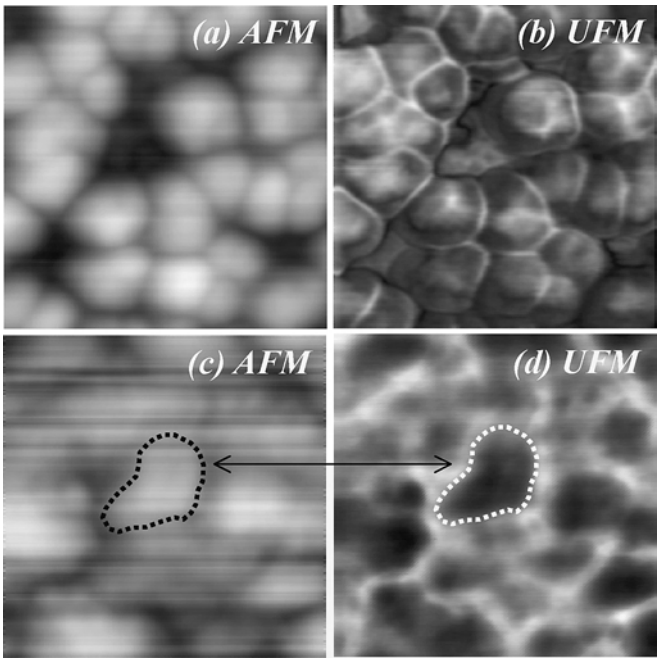


Figure 11. 0.4 μm $\times$ 0.4 μm AFM and UFM scans of Cu grains nucleated on a TaN substrate. (a) AFM and (b) UFM of a 25 nm thick (nominal) Cu film. The UFM contrast is dominated by variations within grains. The bright regions at grain boundaries result from an increased tip/sample contact area. (c) AFM and (d) UFM scans of a 5 nm thick (nominal) Cu film. In the UFM image, the Cu grain appears as a dark area on a bright substrate. The thickness of the outlined grain in (c) and (d) is approximately 5 nm. The medium contrast regions above this grain correspond to Cu regions ~1.5 nm in thickness.

Cu nanograins on a TaN substrate (images have undergone no postimage processing, including flattening). The UFM image clearly separates Cu nanograins from the substrate. A 5 nm thick grain highlighted in Figure 11 has a clear contrast differential from the TaN substrate [53]. From topographic analysis, the medium-contrast areas correspond to Cu as thin as 1.5 nm (~five monolayers). Larger Cu grains from thicker (~25 nm films) are also shown in Figure 11. Here, the entire surface has been covered by Cu, and the UFM contrast largely results from variation of tip/surface contact area at grain boundaries or slight variations of contact stiffness within individual grains.

Although Cu can be considered soft compared to the Si_3N_4 tips typically used for UFM, such may not be the case for materials like TaN, as described above. The limits of applicability for UFM have been examined in detail for such high-stiffness materials. Dinelli et al. considered modulus extraction from UFM analysis of polished silicon and sapphire surfaces [35]. Using the JKRS model to carry out simulations of the tip–sample force–displacement response curves acquired from UFM scans, extracted moduli were in good agreement with expectations if the tip–surface contact area remained relatively constant. This work specifically notes complexities introduced into the UFM force–displacement response curve analysis of the tip radius of curvature variation as a function of applied force. This variation is above and beyond the simple elastic deformation predicted by a simple Hertzian model. For example, Dinelli et al. note that, for investigations of the contact stiffness of silicon, the effective tip radius of curvature may increase from an effective value of 6 nm for a low-load-force (~70 nN) to an effective value of 25 nm for a load force of 280 nN. Such challenges are not unexpected for tip and sample moduli of the same order.

However, despite system-dependent complexities in quantitative modulus extraction, the relative modulus sensitivity of UFM is well established. Kolosov and co-workers examined an impressively wide range of material systems via UFM to explore the capabilities of the technique [39]. In addition to the Ge QD work referenced above, UFM investigations of carbon fiber composites, polymer–polymer composites, two-phase Langmuir–Blodgett films, polymer–glass laminates, and immobilized proteins have been undertaken. In each case, UFM imaging revealed nanoscale mechanical variations associated with nanoscale composition. In the case of thin-film laminates, UFM revealed areas of delamination between the thin film and the substrate. This was demonstrated for thin silicon oxide films deposited on PET (polyethylene terephthalate) polymer substrates [54]. Similar phenomena have been observed very recently for the delamination of metal films on polymers. Figure 12 shows the stress-induced delamination of a 50 nm thick Ta film on a polyaromatic polymer substrate (SiLK® polymer, Dow Chemical Corporation). The topography scan clearly reveals a large hillock feature associated with buckling of the Ta film. Corresponding UFM measurements reveal a clear drop in contrast at the edge of the raised region, denoting a dramatic increase in mechanical compliance of the Ta film. Since no compositional variation of the Ta film was evident via spectroscopic analysis, the compliance loss was assumed to result from delamination. An interesting feature of this UFM scan is the high-contrast region at the crest of the raised Ta buckle. It is possible that the UFM tip is sampling a local increase in contact stiffness due to a micromechanical

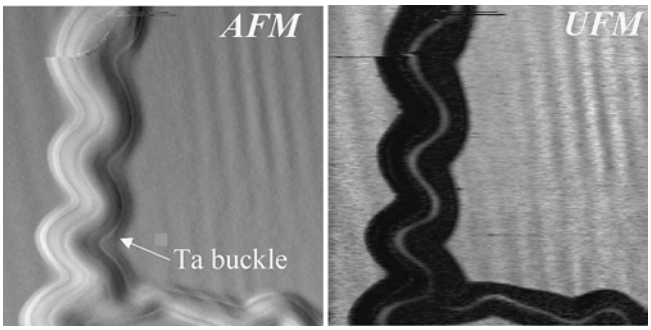


Figure 12. 20 μm $\times$ 20 μm AFM (left) and UFM (right) scans of a 50 nm Ta film deposited on Dow Chemical's SiLK polymer. The AFM scan highlights a buckled section of the Ta film. The corresponding UFM scan reveals the dramatically increased compliance of this region. The high-contrast line running along the center of the delaminated region may result from a micromechanical "arch" effect. The striations at the upper right quadrant of both images correspond to topographic and mechanical compliance corrugations, respectively, presaging the onset of delamination.

arch effect of the Ta film. Investigations into this phenomena are continuing.

As the above discussion notes, surface contact stiffness is affected by a wide array of phenomena related to both the material under investigation as well as the particulars of the nanoprobe used. Consequently, a lower estimate of the sensitivity of UFM to pure modulus variation is of significant interest. Dinelli and co-workers carried out such an investigation using cross-sectional UFM imaging of a GaAs–$Al_xGa_{1-x}As$ supper lattice (lattice constant ~200 nm) [35]. Their work showed that relative changes in the elastic modulus as low as 0.08% (corresponding to a Δx of Al as low as 0.1) were observable in UFM. This is a truly remarkable sensitivity.

Geer and co-workers have recently investigated the use of UFM for integrated-circuit (IC) and nanomaterial characterization [54–57]. The surface mechanical variation of IC interconnect test structures consisting of inlaid (damascene) Al lines in a low-dielectric-constant polymer (benzocyclobutene) was investigated with respect to variations in the polymer modulus resulting from IC processing. The damascene technique used in IC processing consists of the formation of recessed trenches and vias in a dielectric layer via reactive plasma processing (so-called reactive ion etching) [58]. In this manner, an IC interconnect pattern can be "transferred" into a dielectric layer. The recessed features are coated with a TiN barrier layer (to inhibit Al diffusion in the polymer), and subsequently filled with Al metal through vapor deposition [59]. The excess Al is removed via a chemical mechanical planarization (CMP) process resulting in an extremely flat inlaid pattern of transistor interconnect wires [60]. This process is repeated until a multilevel metal interconnect circuit is fabricated that completes the functioning IC circuit. Polymeric dielectrics have been under intense investigation in IC fabrication recently, owing to their low dielectric constant and associated reduction in *RC* losses during chip operation. However, the dramatic mismatch of mechanical properties between metal lines (in some cases <100 nm wide) and polymer dielectrics places limitations on the internal stresses that can be tolerated during the thermal excursions of chip operation. Consequently, to accurately model chip stability, it is important to know the variation in mechanical properties of such dielectrics during their deposition. The reactive ion etch process that is used to "dig" the trenches in the dielectric can also result in changing local polymeric reaction properties [61]. UFM analysis of these structures (Fig. 13) revealed that the reactive ion etch effectively "hardened" the polymer near the exposed surface [54]. This hardening resulted from selective removal by the ion etch of carbon, leaving an SiO_2-rich polymer at the surface of the recessed feature [62]. No such characterization currently used in IC processing has been able to provide such direct evidence of process-induced mechanical variation of materials.

Moreover, UFM has also shown the potential for imaging the surface damage of polymeric materials. Figure 14 shows UFM and topography scans for a region of an Al/polymer IC test structure after CMP [63]. In the UFM image, the lateral lines are Al-filled trenches in the polymer matrix. There is an associated topography (Al lines are slightly depressed relative to the polymer). The UFM also reveals striations

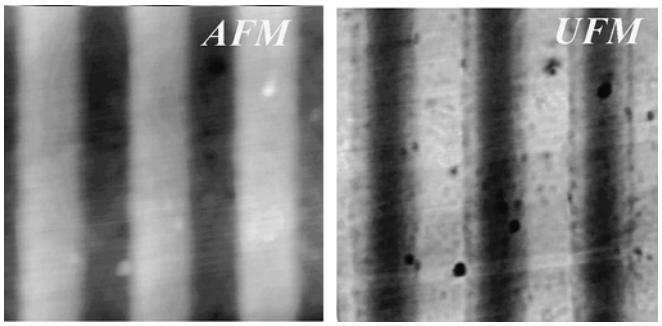


Figure 13. 2.0 μm × 2.0 μm AFM (left) and UFM (right) plan-view scans Al lines (dark regions in topography) in a polymer matrix (bright regions in topography). In the UFM scan, the contrast is reversed, as expected for an elastic image map. Note the local variation of the contrast within the polymer at the edges of the trench walls. This increase in elastic modulus result from silicon–oxide polymer formation in the polymer after exposure to a reactive-ion etch plasma.

crisscrossing the trench field. These striations look similar to scratches, not unexpected after CMP processing. However, the "scratches" have no corresponding feature in the topography image, and only occur in the polymer region. Preliminary investigations associate these features in the UFM image with mechanical stressing in the polymer that results in cross-link bond breakage in the polymer, resulting in a lower local Young's modulus [64]. The reduction of cross linking in these areas also reduces the local shear modulus, and permits relaxation of topographical variations, leading to the featureless topography scan.

Similar behavior relating to the nanomechanical imaging of surface-deformed polymers was carried out by Iwata and co-workers [65]. An AFM tip was used to locally deform or scratch the surface of a polycarbonate film. The local stresses induced the organization of the local carbon nanobundles [66]. In this work, UFM demonstrated

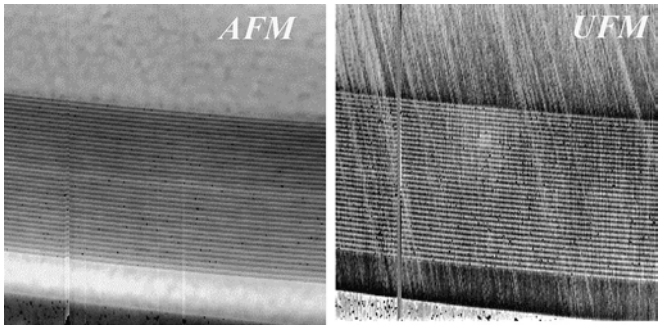


Figure 14. 40 μm × 40 μm AFM (left) and UFM (right) plan-view scans of an Al trench field (horizontal lines) in a polymer matrix. The sample was planarized via chemical mechanical planarization. In the UFM, vertically oriented striations are evident. These striations are not evident in the AFM image, and are believed to result from local bond breaking in the polymer due to the action of the CMP abrasive. Although not visible at this magnification, the striations are not continuous across the Al regions.

sufficient sensitivity for the observation of slight elastic variations of the modified PC surface. The bundles grew in vertical extent with repeated scan scratching, yet the associated bundle mechanical response decreased. The associated contrast of the UFM images of individual bundles decreased, consistent with the threshold behavior of the cantilever response. It was suggested by the authors that the increase of the mechanical bundle compliance is indicative of the microcracks or voids due to the local tip-induced stresses associated with repeated scratching.

Recently, Kolosov has expanded the applicability of UFM by demonstrating so-called waveguide UFM or W-UFM. In such a mode, the ultrasonic vibration is introduced to the cantilever, and not the sample. If a higher order flexural vibration is excited, the tip of the sample will oscillate as an antinode for the standard vibrating beam problem. However, the suppression of the fundamental resonance mode will still restrict the overall cantilever rigidity, resulting in the indentation of the sample against the tip, similar to conventional UFM. In fact, the tip–sample force–displacement response curves for the two techniques are quite similar [67].

Ultrasonic force microscopy offers exceptional sensitivity to surface contact stiffness. Although no quantitative comparison has been carried out, preliminary work by Shekhawat et al. implies that it may be superior to SLAM from the perspective of overall sensitivity. However, the difficulty of quantitative analysis for high-stiffness materials via UFM resulting from contact area variations demands further technique development. A parallel technique developed for nanomechanical imaging offers a different route for the qualitative and quantitative analysis of surface contact stiffness and modulus, so-called acoustic atomic force microscopy (AAFM). Early use of AAFM was reported by Arnold and co-workers [68–70]. It used a slightly different experimental apparatus for mapping the surface elastic response. However, for all intents and purposes, it is essentially the same as UFM when the sample oscillation amplitude is large enough to probe the nonlinear region of the tip–sample force–displacement response curve. The notable difference is the utilization of AAFM in the *linear regime* for extraction of the contact stiffness. Linear regime AAFM (also referred to as ultrasonic AFM by Quate and co-workers [71]; the tendency of each group to define a unique acronym for their particular experimental configuration is confusing for newcomers to the field) is similar to Burnham's SLAM. However, Burnham modeled the AFM cantilever tip as a simple spring, and hence one resonance. The groups of both Arnold and Quate took a more involved approach, and solved the full dynamical equation of a cantilevered tip in contact with the surface. As the surface load of the tip is varied, the resonant frequency spectrum of the cantilever is modified. Hence, the varying surface contact stiffness probed by a cantilevered tip as it is scanned across a sample surface will result in a modified resonance spectra for the cantilever. At a given point on the sample surface, determination of the shift in the cantilever resonance is used to extract the contact stiffness and determine, if desired, the Young's modulus of the surface.

Assuming a linear response, the relation between the surface contact stiffness k_i and a resonant frequency spectra $k_n^2 = f_n\sqrt{4\pi^2 12\rho/(b^2E)}$ [72], the surface-modified cantilever frequency spectrum is given by [69]

$$\frac{k_i}{k_c}\begin{bmatrix} -(\cosh k_nL_1 \sin k_nL_1 - \sinh k_nL_1 \cos k_nL_1) \\ \quad \times (1 + \cos k_nL' \cosh k_nL') \\ +(\cosh k_nL' \sin k_nL' - \sinh k_nL' \cos k_nL') \\ \quad \times (1 - \cos k_nL_1 \cosh k_nL_1) \end{bmatrix} = 2\frac{(k_nL_1)^3}{3}(1 + \cos k_nL \cosh k_nL) \tag{13}$$

Here, L is the total cantilever length, b is the cantilever thickness, L_1 is the base-to-tip length of the cantilever, L' is the tip-to-end length of the cantilever, and E is the Young's modulus of the cantilever material. For cantilever width a, the cantilever spring constant k_c is given by

$$k_c = \frac{Eb^3a}{4L_1^3} \tag{14}$$

Rabe and co-workers used the shift in resonance to qualitatively map the elastic response of a surface [69]. For a given point on a sample, the contact stiffness was approximated. Crozier et al. [71] and Yaralioglu et al. [73] extended this analysis with the development of an iterative algorithm to fit the shifts of multiple resonant peaks of the cantilever spectrum at points on the sample to determine, with high precision, the calibrated contact stiffness. Application of this technique to the film thickness measurement represented an impressive achievement for nanomechanical characterization.

The culmination of this approach for ultrasonic-based scanning probe microscopy has recently been reported by Yamanaka's group [74]. They have modified the conventional AAFM configuration to include a feedback circuit to maintain cantilever resonance as the tip is scanned across the sample with a minimal increase in image acquisition time. Hence, a true nanoscale "map" of the cantilever resonance frequency can be obtained and used for conversion to contact stiffness data (although quantitative conversion has yet to be done). That technique demonstrated some of the highest resolution images to date regarding subsurface dislocations in graphite, and promises a bright future for the direct nanoelastic imaging of materials.

A variety of other related techniques for nanomechanical imaging have been reported; however, the majority represent variations of those discussed with respect to the introduction of lateral oscillation or combinations of high and low frequencies to the tips and samples [75–79]. A unique, and possibly powerful, technique for probing the elastic AND viscoelastic properties of materials at the ultrasonic time scale was recently put forth by Kolosov and Briggs [80–81]. This technique combines W-UFM and UFM. The heterodyne interference of the two vibrations produces a low-frequency amplitude modulation that traverses the nonlinear portion of the force response curve in much the same fashion as the externally modulated ultrasonic vibration in conventional UFM. However, the phase coherence of the two vibrations yields a temporally calibrated phase image that can resolve time shifts on the nanosecond time scale.

Hence, viscoelastic (dynamic) nanomechanical imaging is possible through so-called heterodyne UFM (HFM). In this

mode, two ultrasonic vibrations are applied to the measuring system. One ultrasonic vibration is applied to the tip, which acts as a mechanical waveguide, while a second ultrasonic vibration is applied to the sample (as in conventional UFM). This results in coherent oscillations of the tip and sample surfaces at the local point of contact [80]

$$\begin{aligned} z_S &= A_S \cos(\omega_S t + \phi_S) \\ z_{\text{tip}} &= A_{\text{tip}} \cos(\omega_{\text{tip}} t) \end{aligned} \tag{15}$$

Here, the tip vibration waveform acts as a temporal reference. The temporal phase delay associated with the surface vibration waveform is designated by ϕ_S. A linear force–displacement relationship between the tip and sample would result in no average cantilever deflection. However, a nonlinear force–displacement curve yields a heterodyne coupling between the two oscillations. This is seen by expanding the tip–sample force in a Taylor's series to second order in the tip–sample separation, and calculating the average force responsible for cantilever deflection

$$\begin{aligned} &k_c \langle z_{\text{tip}}(\omega_{\text{tip}}, \omega_S, A_{\text{tip}} A_S) \rangle \\ &\quad = \frac{1}{\tau} \int_0^\tau F(z_{\text{tip}} - z_S)\, dt \\ &\quad = \frac{1}{\tau} \int_0^\tau \left(\chi_1 (z_{\text{tip}} - z_S) + \chi_2 (z_{\text{tip}} - z_S)^2 + \cdots \right) dt \end{aligned} \tag{16}$$

χ_1 and χ_2 denote first- and second-order "spring constants," respectively, associated with the force–displacement curve. The cantilever deflections resulting from both oscillations [Eq. (15)] are substituted into Eq. (16) and time averaged over the response time τ of the SPM photodiode detector (hundreds of kilohertz). Only the nonlinear term is nonvanishing. Likewise, the high-frequency terms are replaced by their rms values, and the remaining time dependence corresponds to the heterodyne or difference frequency term

$$\begin{aligned} &\langle z_{\text{tip}}(\omega_{\text{tip}}, \omega_S, A_{\text{tip}P} A_S) \rangle \\ &\quad = \frac{\chi_2}{k_c} \left\{ \frac{A_{\text{tip}}^2}{2} - A_{\text{tip}} A_S \cos[(\omega_{\text{tip}P} - \omega_S)t - \phi_S] + \frac{A_S^2}{2} \right\} \end{aligned} \tag{17}$$

The deflection of the tip is now dynamically linked to the viscoelastic response of the sample to the ultrasonic vibration. The phase term ϕ_S represents the dissipative lag/lead in the surface response with respect to the tip reference frequency. In the vicinity of an interface, this dissipation is directly linked to a local loss tangent, indicative of the adhesive or interfacial bonding strength. For continuous materials, the dissipation is due to a classical viscosity. Extracting the spatial dependence of this phase term provides image contrast indicative of the dynamics of materials, material interfaces, and defect structures with the same nanometer spatial resolution as UFM.

Experimentally, the two oscillations are applied to the tip and sample by two matched piezocrystals attached to the tip substrate and the base of the sample, respectively, and driven by a separate waveform with a mixer/filter circuit providing the heterodyne frequency to an RF lock-in amplifier for amplitude (χ_2) and phase (ϕ_S) extraction. Although the amplitude calibration is more complex than that for UFM, the phase can be imaged directly, as displayed in Figure 15 for an integrated-circuit test structure consisting of square 5 μm polymer pads inlaid within an Al layer [60, 82]. The right panel data of Figure 15 were acquired with both ultrasonic vibration amplitudes above threshold for Al and the polymer. The left panel data of Figure 15 were acquired with both amplitudes below threshold. No phase variation is seen in the latter since the response is linear, validating the simple model presented above. The phase difference between the Al and polymer regions ~10 ns. Data were taken at a carrier frequency of 3 MHz. Recent efforts in our laboratory have demonstrated HFM operation up to 80 MHz. The 1° relative phase resolution of our HFM system provides a temporal (phase lag) resolution less than 40 ps.

3. NANOTUBE MATERIALS

3.1. Introduction: Description of Molecular Nanotubes

In 1991, nanotubes of carbon were discovered by Sumio Iijima in carbon deposits on an arc-deposition system similar to that used for the production of C_{60} (termed Buckminsterfullerene) [3]. From a structural perspective, a single-walled carbon nanotube (SWNT) can be visualized as consisting of a single (graphene) sheet rolled into a tube. To maintain the periodic structure of the 2-D honeycomb lattice formed by the carbon atoms within the sheet, there is a discrete relationship between the tube diameter and the chirality of the tube [83, 84]. This relationship is summarized in Figure 16, and utilizes a simple 2-D lattice notation. The diameter and chirality of a given SWNT are defined by an integer pair (n, m) which denotes a point on the edge of a graphene sheet that is mapped onto the origin (0, 0) as the sheet is rolled. Rolled (n, 0) graphene sheets are referred to as zigzag tubes as C atoms on an open end form such a pattern. Rolled (n, n) graphene sheets are referred to as armchair tubes as the C atoms on an open end form undulating hexagonal tiles (see Fig. 16). Rolled (n, m) sheets, where $n \neq m$, are typically referred to as chiral, although the use of this taxonomy varies throughout the literature.

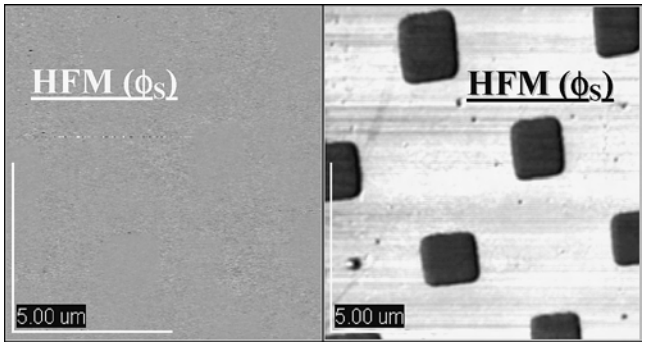


Figure 15. HFM phase maps (10 μm × 10 μm scans) for tip and sample vibrations below (left) and above (right) threshold amplitude, respectively, for polymer/Al test structure. The dark regions are polymeric plugs in an Al field. The structure was fabricated using IC processing. The reduced contrast for the polymeric regions correspond to an increase in the viscoelasticity compared to the Al.

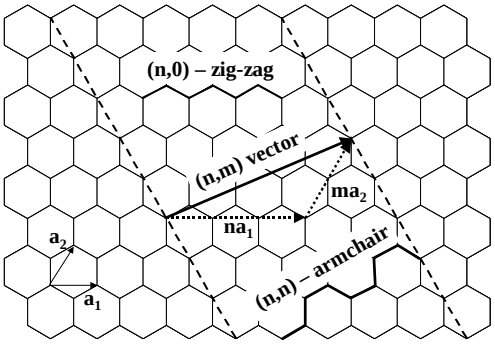


Figure 16. Schematic diagram outlining how a 2-D graphite sheet is "rolled" to form a nanotube. The diagonal dotted lines denote graphene sheet edges that, if joined, would form a (3, 2) nanotube.

Multiwall carbon nanotubes (MWNTs) consist of multiple nested (graphitic) sheets rolled into a tube. The spacing of these sheets has been experimentally measured at 0.34 nm, only slightly larger (3–5%) than the c spacing of conventional graphite [84]. SWNTs typically possess diameters in the 1–5 nm range, and lengths ranging to many microns. Depending upon the method of synthesis and purification, SWNTs entangle to form nanoropes [85–86]. The cross section of these ropes consists of individual SWNTs on a triangular lattice. For example, the SWNT ropes investigated by Yu et al. consisted of 1.35 diameter SWNTs on a triangular lattice with a center-to-center distance of 1.7 nm [87]. As expected, MWNT diameters are significantly larger, ranging from 7 to over 200 nm. Multiwall NTs also form fibrous ropes, the mechanical properties of which will be discussed next.

From a mechanics perspective, two fundamental properties of NTs are of interest. First, the planar sp^2 bonding of the carbon atoms in the graphene sheets provides for an enormous mechanical modulus, that is, a very low strain is seen for a given in-plane applied stress [88]. And, in stark contrast to fourfold coordinated sp^3 bonding of carbon in diamond or diamond-like materials, the quantum-mechanical electron energy levels of an sp^2-bonded graphene sheet may permit electron conduction [89–90]. Second, the energetics of NT growth are often suitable for the production of crystal defect-free SWNTs or MWNTs [91]. This is a truly remarkable mechanical feature in that the presence of crystalline defects dramatically impacts mechanical properties.

3.2. Mechanical Definitions, Theoretical Predictions for NT Mechanical Properties, and Experimental Studies

As noted by Salvetat et al., [92], knowledge of the Young's modulus of a material is the first step toward its use in practical structural applications. For a long, thin rod of arbitrary cross-sectional area A_o, the Young's modulus is defined as the ratio of the axially applied stress (F/A_o) to the fractional axial deformation $(\delta l/l_o)$: $E = (Fl_o)/(\delta l/A_o)$. The Young's modulus can also be defined in terms of the bulk modulus K and the shear modulus μ of a material: $E = 9K\mu/(3K + \mu)$. In the simplest terms, the shear modulus is the ratio of the applied stress to the strain of a material in the case where a deformation changes the shape of a body, but not its volume. The opposite case is true for the bulk modulus. It represents the stress/strain ratio for a deformation changing the volume of a body, but not its shape (hydrostatic). Both K and μ are defined assuming a linear relationship between the applied stress and strain. The Young's modulus for both SWNTs and MWNTs has been calculated, simulated, and measured in a variety of methods which are described in more detail below. Some groups model nanotubes as deformed sheets wherein the strain energy due to bending is linearly proportional to a bending modulus κ, defined as the second derivative of the surface energy density with respect to curvature. For a tube, $\kappa = Ea/(1 - \nu^2)$, where ν is Poisson's ratio and a is the tube wall thickness. Also of interest for the nanomechanics of NTs is the tensile or yield strength (defined as the stress beyond which the strain varies nonlinearly or, alternately, as the maximum stress which may be applied to the material without perturbing its stability [93]).

One of the earliest theoretical studies of the mechanical properties of SWNTs was carried out by Robertson et al. [83]. They investigated the lattice energetic of carbon NTs for all (n, m) nanotubes with radii less than 0.9 nm using two related many-body empirical potentials and local density functional (LDF) theory [94]. The many-body potentials included atomic core–core repulsive terms and attractive terms related to the C valence electrons. It was found that the tensile or compressive strain energy per carbon atom of an SWNT relative to an unstrained tube varied as $1/R^2$, where R is the tube radius. In addition, it was found that the Young's modulus associated with tube elongation decreased with decreasing radius, that is, smaller nanotubes became softer, depending on the helical (n, m) structure of the nanotube. The former observation agrees well with the expectation of continuum elastic models for the strain energy density of a uniform tube [83, 95]

$$\frac{\sigma}{A} = \frac{Ea^3}{24R^2} \tag{18}$$

where A is the area of the NT, a is the representative thickness of the NT wall (~0.34 nm), and E is the Young's modulus. Robertson noted that agreement between simulation and continuum elastic theory persisted to radii as low as 0.2 nm. This result was not completely surprising. A tube consisting of a graphene sheet would be expected to maintain a Young's modulus near that of the C_{11} elastic constant in graphite (corresponding to the elastic modulus parallel to the basal plane), as long as the atomic level deformation induced by the wrapping was low. This is obviously the case for large radii.

In studying the deformation properties of carbon SWNTs, Bernholc and co-workers calculated the Young's modulus and Poisson's ratio via a molecular dynamics framework based on many-body interatomic potentials [96–98]. As with the earlier work of Robertson et al. [83], a high degree of correspondence between the modeled mechanical response and results of classical mechanical theory was found. Their MD simulations predicted a modulus of 5.5 TPa and a Poisson's ratio of 0.19 [96]. The latter is fully 550% larger

than the C_{11} elastic constant (basal plane modulus) of graphite, while the Poisson's ratio is significantly lower than the value of 0.33–0.34 of graphite [50]. At this point, a comment is required regarding the calculation of E for atomically thin tubular shells. From the perspective of continuum elastic energetics, the Young's modulus is also defined as the second derivative of the system energy with respect to strain at equilibrium normalized to system volume. This is the practical definition used for virtually all simulations or analytical models as the energy functional is directly accessible. However, it opens the question of the actual wall thickness of an SWNT. At molecular-length scales, it may not be reasonable to characterize the thickness of an SWNT with the lattice spacing of MWNTs or the c-axis spacing of graphite. Consequently, the use of an elastic modulus such as E, which is, by definition, normalized to a classically defined cross-sectional area, may have little meaning in that a well-defined value for the NT cross-sectional area does not exist. More recently, several groups working on this problem from a simulation or theoretical perspective have stated results only in terms of the product aE to avoid the ambiguity inherent in applying a classical "wall thickness" to a structure of single-atom thickness [99–100]. For the MD simulations previously described, a wall thickness of 0.066 nm was used. When replaced with the value of 0.34 nm (an upper estimate for the wall thickness in MWNTs) and akin to the c-axis spacing in graphite, a value for E of 1.07 TPa is obtained. Suffice it to say that, when compared in terms of aE, the MD work presented here continues to agree with more recent theoretical (and, as will be shown, experimental) results.

An empirical force-constant model, wherein the atomic interactions near an SWNT equilibrium structure are approximated by the sum of pairwise harmonic potentials, was used by Lu to calculate elastic properties of SWNTs and nanoropes [101]. Values of $E = 0.97$ TPa were obtained using an SWNT wall thickness of 0.34 nm, and were virtually constant as a function of SWNT helicity and radius. A value for Poisson's ratio of 0.28 was determined that possessed the same insensitivity to (n, m). The Young's moduli of crystalline nanoropes of SWNTs were calculated, and ranged from 0.80 to 0.43 TPa as the number tubes per rope varied from 5 to 15, respectively.

More recently, a broad range of computational and analytical approaches has been applied to the accurate calculation of Young's modulus in CNTs. Yao and Lordi took a finite temperature, molecular dynamics (MD) simulation approach for the calculation of E for SWNTs [102]. Using the universal force field of Rappé and co-workers [17], simulations were carried out on SWNTs 10 nm in length, ranging in diameter from approximately 0.2 nm for a (5, 5) tube to 1.4 nm for a (20, 20) tube. For relatively large (20, 20) tubes, the Young's modulus determined from the simulations was approximately 0.97 TPa. This is quite close to the value of C_{11} measured in graphite ($\sim$1 TPa [104]). However, as the SWNT radius was decreased, the MD simulations predicted a 15% increase in E to 1.12 TPa for the (5, 5) SWNT. The increase was attributed in large part to corresponding increases in torsional strain that develop as the C sp^2 bonds are deformed at a smaller tube radius. The variation of strain energy with NT size had been previously studied [105–106], but the ability for the MD simulation to potentially predict this behavior more accurately was attributed directly to the inclusion of a torsional term in the bond energy calculation, which turns out to be most sensitive to the variation of E with decreasing radius. Yao and Lordi also simulated thermal vibrations of SWNTs clamped at one end, similar to TEM-based techniques, to experimentally measure E (see below). Their simulations, ranging from temperatures of 300 to 2000 K, yielded temperature-independent vibrational mode frequencies that matched well with continuum elastic model predictions.

Similar studies were carried out by Goddard and co-workers using molecular dynamics and molecular mechanics techniques, but for (n, n), $(n, 0)$, and $(2n, n)$ SWNTs [107]. They obtained values for the bending modulus κ of 0.96 TPa for (n, n) tubes, 0.91 TPa for $(n, 0)$ tubes, and 0.94 TPa for $(2n, n)$ tubes. They also calculated Young's moduli for SWNT bundles or ropes. Within the bundle, the SWNT axes define a triangular lattice. The lattice spacing varied from 1.68 nm for (n, n) tube bundles to 1.65 nm for $(n, 0)$ and $(2n, n)$ tube bundles. Strain energies for these ropes were calculated, and Young's moduli of 640, 648, and 673 GPa for the (n, n), $(n, 0)$, and $(2n, n)$ ropes, respectively, were obtained. The comparison between the bending moduli calculated in this work and Young's moduli of other techniques is complicated by a dimensional inconsistency. Although the strain energies are defined conventionally, Goddard and co-workers report the bending moduli in units of gigapascals, and do not define the normalization of the wall thickness required for conventional definitions of κ [107].

Rubio and co-workers have carried out predictive calculations for E of SWNTs using tight-binding (TB) and pseudopotential-based density function theory methodologies [99]. To avoid the wall thickness problem, the Young's modulus, alternately defined in elastic systems as the volume-normalized second derivative of the system energy with respect to strain (at equilibrium), is instead normalized to the SWNT surface area. Their TB model followed that of Porezag et al. [108], containing band-energy terms and a repulsive pair potential. As with similar techniques used for band calculations of conventional materials [109], the successful application of this technique depends upon a generalized eigenvalue equation for a nonorthogonal basis set. As with previous continuum elastic and calculational approaches, an inverse quadratic dependence of strain energy on tube diameter was found [83]. Assuming a carbon SWNT wall thickness of 0.34 nm, their corresponding results for E ranged from 1.26 TPa for a (20, 0) carbon SWNT to 1.22 TPa for a (6, 6) SWNT. In addition, the authors calculated the Poisson's ratio, defined here as

$$\nu \cdot \varepsilon = -\frac{R - R_{\text{eq}}}{R_{\text{eq}}} \tag{19}$$

The Poisson's ratio varied from 0.27 for a (20, 0) tube to 0.25 for a (6, 6) tube. Considering the sidewall thickness used for the determination of E, these results agreed reasonably well with other approaches, and confirmed that, in certain respects, the elastic behavior of carbon SWNTs should mirror that of in-plane graphite. An important aspect of this

work is the prediction that E will decrease rapidly as the tube radius decreases. This tends to agree with earlier comments made by Robertson et al. [83], but contradicts the intuitively appealing result of other groups that E should increase with smaller radius tubes due to the increased energy associated with the rolling strain [102]. Despite the significant theoretical and modeling efforts ongoing for nanotube structures, this issue seems not to be resolved.

Mechanical strain energies and Young's modulus for straight and bent SWNTs were also calculated by Xin et al. [110]. In contrast to prior investigations that used continuum elastic theory, molecular dynamics simulations, or empirical many-body potentials, a band-theory approach was taken here [111–112]. It was noted from a straightforward calculation that the strain energy of straight SWNTs originates from the curvature-induced electronic energy change. Consequently, determination of the electronic energy of all of the occupied bands was carried out to provide predictions for SWNT mechanical properties. The band-structure calculation utilized a nearest neighbor tight-binding approach [111]. Such approaches have been widely used to calculate graphene and SWNT electrical properties. Strain energies were calculated for a tube of cross-sectional radius R subjected to a bending deformation (bending radius $= \rho$). The R dependence of the bending energy E_b related to the strain energy E_s of a straight tube determined from the tight-binding model agreed well with continuum elastic theory predictions [113]

$$E_b = E_s + \frac{\lambda}{R^2} \tag{20}$$

From the tight-binding calculation carried out by Xin and co-workers, $\lambda = 0.7$ eV/atom. This is a further confirmation that continuum elastic theory provides a valid framework within which to predict nanotube mechanical behavior.

In this same work, however, a Young's modulus of approximately 5 TPa is predicted. This magnitude exceeds other estimates by a factor of 5. However, it should be noted that the SWNT wall thickness calculated by Xin et al. is approximately 0.07 nm, and differs significantly from the conventional value of $a = 0.34$ nm used for the calculation of E by most other authors (noted previously). In terms of the product aE, the results from this tight-binding model match well with other simulations.

Theoretical calculations of SWNT mechanical properties have also been carried out using local density approximation cluster models. As applied by Zhou et al., a linear combination of atomic orbitals is used to solve the Schroedinger equation directly, applying appropriate boundary conditions to remedy the problem of dangling bonds present at the edges of the cluster [93]. Such an approach makes fewer initial restrictions on the form of the electron wave function as do tight-binding approaches. A 156 atom NT was used for purposes of calculation [114–115]. A value of Young's modulus of 0.76 TPa was calculated, somewhat lower than previously discussed values (assuming a wall thickness of 0.34 nm). A value of Poisson's ratio of 0.32 (slightly lower than that of graphite, ~0.33–0.34) was also determined from these calculations. It should be noted that, unlike molecular dynamics approaches, the density-of-states methodology used here also provides specific information regarding carbon–carbon bonding states. It was found that, upon rolling the graphene sheet into an (n, n) conformation, the overlap integral of σ bonds in the graphene plane decreased by approximately 7%. In contrast, the overlap integral of π bonds increased by about 2.4%, and the total overlap integral decreased by approximately 4%. The decrease of the overall overlap integral leads to a lower net bonding energy compared to graphite. In addition, the reduced bond overlap will broaden the effective interatomic electrostatic potential well, and reduce the Young's modulus compared with that of basal plane graphite (~1 TPa). It is interesting to note that, in this work, the curvature effect tends to weaken the σ bonding, but strengthens the π bonding, which may have implications with respect to the electronic functionality of NTs.

Fewer theoretical and simulation efforts have been carried out with respect to the elastic moduli of MWNTs. Using the empirical force constant model referenced earlier, Lu predicted an increase of E with the number of walls from a value of 0.97 TPa for a single (5, 5) SWNT to a value of 1.11 TPa for an MWNT with seven nested tubes about a (5, 5) core [101]. A Poisson's ratio value of 0.27 was insensitive to the number of MWNT walls. For these studies, as most others, a 0.34 nm spacing is assumed. The critical aspect of MWNT modeling is proper inclusion of the electrostatic energies associated with the adjacent graphitic walls. No dramatic effects have been predicted with respect to linear response to elastic strain. However, several groups point to significant effects of wall–wall coupling in nonlinear (inelastic) deformation, and they will be discussed next.

Experimental determination of the Young's modulus for SWNTs and MWNTs has utilized a wide array of approaches. The groundbreaking work was carried out on MWNTs by Treacy et al., resulting from an observation that the free ends of cantilevered SWNTs appeared fuzzy in a TEM [116]. Subsequent thermal studies revealed that poor TEM focus resulted from thermal vibrations. Using classical mechanical beam theory, the vibration frequency spectrum ω_n of a uniform clamped tube [116, 95]

$$\omega_n = \frac{\beta_n^2}{2L^2}\sqrt{\frac{E\left(R_i^2 + R_o^2\right)}{\rho}} \tag{21}$$

was used to estimate the E experimentally, assuming a thermodynamic equipartition concerning the SWNT vibrational mode energies. Above, R_i and R_o represent the inner and outer SWNT tube radii, respectively, ρ is the in-plane mass density, L is the tube length, and β_n is a set of numerical coefficients. Fitting the observed thermal dependence of the vibration amplitude to the above expression, E_{measured} varied from 0.4 to 3.7 TPa for wall numbers ranging from approximately 5 to >25. Although the uncertainties in the data were quite large and dominated by experimental errors, there was an increase in E attributed to thinner tubes. These studies were extended and refined by Krishnan et al. with respect to SWNTs [117]. Using the same basic experimental techniques and stochastically driven oscillator theory for the SWNTs, a refined error analysis was presented, and an

average Young's modulus of 1.3–0.4/+0.6 TPa was determined. Similar to the work on MWNTs, measured values of E for SWNTs ranged from ~0.3 to 2.7 TPa. The origin of experimental errors stems from the accurate determination of the thermal vibration amplitude to the accurate calculation of the free length of individual NTs. Regardless, these studies represent a breakthrough for nanomechanical measurement. Note that the clamped beam models used to calculate E from experimental data require similar assumptions regarding the NT wall thickness as did the theoretical efforts [95]. All experimental work summarized here typically used a value of 0.34 nm SWNT wall thickness (or MWNT wall spacing). Hence, their elasticity data should be compared with correspondingly normalized values of E from theoretical calculations.

More direct measurements of SWNT mechanical response were undertaken by Salvetat et al. that exploited the surface adsorption of SWNT ropes to nanoporous templates to form suspended SWNT rope "beams" clamped at both ends by surface van der Waal forces [118]. The ropes were modeled as a continuous cylindrical object characterized by an elastic modulus E_r (Young's modulus renormalized by bending) and a shear modulus μ. The elastic modulus decreased from a maximum of 1.2 TPa (for a 3 nm diameter rope) to values as low as 0.07 TPa for a 20 nm diameter SWNT rope. The measured shear modulus acted in many respects similarly, ranging from 6.5 to 0.7 GPa for ropes of diameter 4.5 and 20 nm, respectively. Although the error values for these elastic constant measurements are reported at $\pm 50\%$, the estimation of E_r for the thinnest ropes is consistent with that of individual NTs. In this work, the shearing was assumed to occur between SWNTs only, that is, shearing of individual SWNTs was ignored. An inherent difficulty in work of this sort is the possibility of defects or breaks of single SWNTs in the ropes. Such effects are cited as a possible source for error, and are primarily responsible for the low value of μ.

In contrast to the beam-bending approach utilized above, Ruoff and co-workers conducted a series of impressive experiments wherein SWNT ropes were literally welded to a cantilevered probe tip inside an SEM, and carried out direct tensile testing driven by a piezoelectric deflector [119]. In these experiments, the base of the nanorope undergoing testing was entangled within a nanorope mat produced via the standard arc-discharge method. By knowledge of the force constant of the cantilevered probe tip, the Young's modulus of multiple nanoropes (assumed to be comprised of (10, 10) SWNTs) was directly measured, and varied from 0.32 to 1.47 TPa, with a mean of 1.0 TPa. These data are in excellent agreement with other experimental measurements and a larger body of theoretical work.

Since the increased size of multiwalled carbon nanotubes and corresponding MWNT ropes increases experimental accessibility, a wide range of mechanical measurements have been performed on these systems. Pan et al. pulled individual MWNT ropes as long as 2 mm to carry out direct tensile strength measurements using conventional test kits [120]. The linear portion of the stress–strain curve thus generated was used to directly measure E. The average value thus determined was 0.45 ± 0.23 TPa. Akita and co-workers used a slightly more elegant approach by attaching an MWNT segment directly to a micromachined cantilever while using a second cantilever applied to the tube to carry out compressive strain testing. The MWNT was considered as a classic beam, and the measured buckling force was assumed described by Euler beam theory ($F_{\text{buckling}} \sim \pi^2 EI/2L^2$; L = length, I = MWNT inertial moment) to extract a Young's modulus of 0.46 TPa (no errors were quoted). Using an experimental approach similar to Treacy and co-workers to extract elastic moduli from nanotubes, Gao considered MWNTs grown pyrolytically [121]. Similar thermal vibrations were observed via TEM from individual MWNTs, and the conventional analogy between the MWNT and a vibrating uniform beam clamped at one end was used to infer the modulus from the thermally stimulated vibration amplitudes [95]. In this work, the authors erroneously label the modulus used in such a model as a bending modulus. For the classical clamped beam problem, the modulus is defined as Young's modulus [4, 95]. Their results suggest an average value of E for their tubes of 0.03 TPa. This value is dramatically lower than prior experiments. The authors note that the expected high density of structural defects in pyrolytically grown MWNTs is the most likely cause. For MWNTs possessing so-called volume defects (areas of the tube with reduced radii), the effective value of E was as low as 0.002 TPa. It is clear that structural defects in NT structures can have dramatic effects on the mechanical properties.

The previous discussions have centered on nanotubes made from carbon. However, there is no physical principle limiting the composition of such structures to carbon. Vaccarini et al. have nicely summarized theoretical and some experimental results for nanotube compositions ranging from the standard carbon to BN, BN_3, BCN_2, GaSe, WS_2, and MoSe. Hernandez and co-workers carried out theoretical investigations of BN, BC_3, BCN_2, and C_3N_4 SWNTs as well [88]. From a physical point of view, all compounds exhibiting layered structures in the bulk phase are likely to form nanotubes, with the caveat that the "layered" structure possesses dramatically anisotropic bonding as is found in graphite. For example, BN carbon nanotubes possess the same honeycomb sheet structure of carbon NTs; however, B and N atoms populate alternating sites. The same (n, n) notation can be used to describe the tube structure, although the "unit hexagon" is composed of equal B and N atoms. Utilizing TB models (see above), the BN SWNT Young's modulus varied from 0.75 to 0.92 TPa for tube radii varying from 0.5 to 2.5 nm, respectively. The increase of modulus with increasing tube radius mirrored theoretical results for corresponding carbon SWNTs using the same computational approach. The same authors note that the Poisson's ratio was similar for all NT compositions studied [88]. And once again, the results for BN nanotubes compared well with simple continuum elastic approaches.

Molecular dynamics and ab initio total energy methods were applied to BN nanotube modeling by Srivastava et al. [122]. A value $E = 1.2$ TPa was determined. Contrary to carbon nanotubes, however, the bond strain inherent in confinement of B–N sheets in tube form tends to expel the N atoms and subduct B atoms, providing an atomic-scale corrugation at the tube surface. Hernandez and co-workers

obtained lower values for BN SWNT Young's moduli generated from TB models [99]. 0.84 TPa $< E <$ 0.91 TPa were determined for tube geometries ranging from (6, 6) to (20, 0). No significant dependence on size or chirality was found. Typically, other composition SWNTs displayed Young's moduli in similar ranges, with the exception of C_3N_4 tubes where E dropped to ~0.6 TPa. In all cases, the highest values of E from simulations corresponded to carbon tubes. From an experimental point of view, measurements of BN nanotube elastic moduli were carried out in precise analogy to those for carbon tubes. Using the thermal vibration amplitude methodology, Chopra and Zettl measured the BN MWNT Young's modulus to be 1.22 ± 0.24 TPa [123]. This compares well with the theoretical result of Srivastava et al. [122]; however, it exceeds estimates from other simulations [93]. Still, the body of experimental and theoretical work on noncarbon nanotubes makes clear the ubiquity of extremely high-modulus nanotube structures provided a high degree of structural perfection exists.

From the viewpoint of elastic properties, an impressive degree of agreement between experimental and theoretical studies of carbon (and other) SWNTs and MWNTs is evident. The question arises as to the promise of nanotechnology. Do the high values of E constitute an emergent nanomechanical property? Although the point may be argued, it is clear that easily synthesized arrays of these nanostructures with near-perfect elastic properties can be produced. The fact that the measurements confirm that these materials possess superior properties compared to larger (microscale) carbon fibers justifies the classification of NTs as nanostructured materials with emergent properties.

The remainder of this section will provide a brief overview of investigations of inelastic properties of nanotubes, namely, tensile strength (defined experimentally as the maximum force sustainable during a stress–strain test) and strain-induced defect generation. Utilizing the tensile strain test kit described earlier, Pan et al. measured the tensile strength of MWNT ropes to be 1.33 GPa [120]. Early theoretical studies [124] suggested that the tensile strength of carbon NTs should be similar to carbon whiskers (~20 GPa). Further refinement of the experimental studies by the same group, including a Weibull analysis of a collection of MWNT ropes, yielded an average tensile strength value of 3.6 GPa [125]. More recently, Yu and co-workers undertook high-resolution tensile strength studies of carbon SWNT ropes, and measured breaking strengths that ranged from 13 to 52 GPa [126]. The experimental mean value was ~30 GPa. The wide variance of the breaking strength values are attributed to various experimental and SWNT rope geometric factors and, most importantly, the lack of information regarding the continuity of individual SWNTs within the rope. However, their results clearly establish that a 5% strain is reproducibly achievable in carbon NT systems.

This value of this so-called breaking strain has been verified by a number of experimental studies as representative of a strain necessary for crystalline defect generation in the graphene or graphitic sheets of SWNTs or MWNTs, respectively. Before reviewing these results, it is important to appreciate the C–C bond orientation as a function of NT helicity. For the so-called armchair (n, n) NT geometry, C–C bonds are either normal to the axis of the tube or rotated by angles of $\pm 30°$ to that axis. Zigzag $(n, 0)$ tubes possess C–C bonds either parallel to the tube axis or rotated by angle of $\pm 60°$ to that axis. The normally oriented bonds of the (n, n) tubes present the greatest moment to axial strain, compressive or tensile. Consequently, inelastic deformation would be expected to vary significantly on the tube helicity. Although the currently available experimental data present only anecdotal corroboration of this observation, theoretical and simulation-based studies confirm, in detail, this effect.

Experimental studies have also demonstrated the incredible flexibility of NTs against nonrecoverable deformation. Falvo et al. reported AFM studies of MWNTs wherein tubes were subjected to large buckling strains by the AFM tip. Kink angles varied from 90 to ~180° [127]. Of particular interest were two phenomena. First, topographic rippling was observed along the tube axis, as is expected from continuum shell theory [128, 129]. Away from the kinks, the rippling was somewhat recoverable; however, topographic variations at the kink positions appeared permanent. In no case was tube breaking observed, a testament to the flexibility and strength of NTs, even in a highly deformed state. In fact, topographical imaging studies have been carried out on completely collapsed MWNTs [126] or nanotube ribbons that possessed remarkable flexibility. It should be noted that an isolated ribbon or collapsed NT is energetically less stable than its tubular analog [107]. The aforementioned experimental analysis explicitly calculated the stabilizing effect of the attractive NT/substrate interaction to demonstrate its crucial role in observing such deformed NT structures.

Nanotube kinking was observed via TEM studies soon after their discovery [84, 130]. Continuum shell analysis was typically invoked to describe these deformations, although no formal justification for the applicability of those theories was provided, with the exception of the $1/R^2$ dependence of the "elastic" nanotube strain energy compared to that of a single graphene sheet. Groundbreaking work regarding a theoretical basis for nanotube deformation was carried out by Bernolc and co-workers through MD simulations and a continuum shell model [96, 130]. The authors considered inelastic deformation formation in carbon SWNTs under compression, bending, and torsion. With respect to bending strain, simulations for a 1.2 nm diameter tube predicted kink formation at a bend angle of ~30°. The critical curvature C_c associated with tube kinking is described by an empirical relation fit to the simulation results

$$C_c = \frac{1.49}{d^2}\left[1 + \frac{9.89}{d^5}\cos(6h)\right] \tag{22}$$

where d is the tube diameter and h characterizes the helicity ranging from 0 [zigzag or $(n, 0)$] to $\pi/6$ [armchair or (n, n)]. The formation of double kinks was observed and, interestingly, simulations implied that the bending deformations were completely recoverable up to bending angles of 110°. Deformation of SWNTs under compressive strain was also predicted by their simulations. For strain $\varepsilon > 5\%$, a characteristic double-fin flattening occurred. At higher strains, a three-fin flattening occurred before broad buckling deformations set in. MD simulations of excessive torsional stress resulted in the flattening and twisting of the nanotube into a ribbon-like structure [91]. Utilizing a continuum shell model,

the authors considered inelastic deformation perturbations in terms of the stability of axial and azimuthal ripples. A stability analysis against such rippling yielded critical strains $\varepsilon_c = (0.077\ \text{nm})d^{-1}$, which agree well with the MD results noted previously.

MD simulations were further used to investigate strain release in SWNTs through defect generation. Bernholc and co-workers [131] investigated the static and dynamical properties of such tubes under uniaxial tension. At high temperatures (~1800 K), (5, 5) tubes under a 10% axial strain spontaneously generated crystal defects via C–C bond rotation. This defect, referred to as a Stone–Wales transformation, rotates the C–C bond normal to the tube axis, generating a double pentagon–heptagon defect (so-called 5–7–7–5 defect). This is pictured in Figure 17. Using classical potential calculations, the formation energy of 5–7–7–5 defect formation was calculated, and shown to be stable for various (n, n) SWNTs at strains between 5 and 6%. If it is assumed that such defect generation is a precursor to mechanical yield, these calculations support the aforementioned experimental studies. From the onset of the 5–7–7–5 defects, simulations predict unbinding of the defect pairs, and subsequent strain-induced motion along appropriate armchair glide planes. Higher order defect generation (e.g., 5–7–5–8–5) were also predicted prior to tube "breaking" which, in simulations, proceeded by unraveling of carbon strips from the edge of the tube until final separation occurred. The motion of defects thus implies a certain degree of ductility for SWNTs.

The propensity of (n, n) SWNTs to generate C–C bond rotation defects has important implications regarding the relative ductile behavior of nanotubes as a function of helicity [132]. In an extension of the aforementioned simulation work, Nardelli examined defect generation in $(n, 0)$ and (n, m) tubes. As expected, the orientation of C–C bonds in $(n, 0)$ tubes dramatically altered the stability of C–C bond rotation defects under applied strains [132]. In fact, the lack of defect generation and subsequent defect motion in these NTs dramatically reduced their ductile properties. Consequently, a theoretical "ductility phase diagram" was developed, and is adapted in Figure 18 to describe the (n, m) domains where brittle versus ductile behavior is expected. Experimentally, confirmation of this phase diagram is challenging due to the uncertainty of temperature effects and

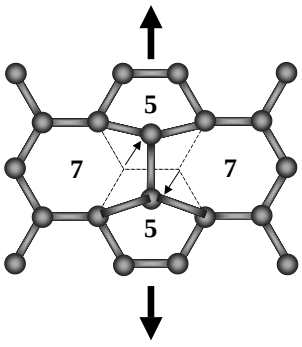


Figure 17. Schematic of a Stone–Wales transformation occuring in an (n, n) nanotube under axial tensile stress. For sufficient stress, the defects can separate and alter the local chirality of the tube.

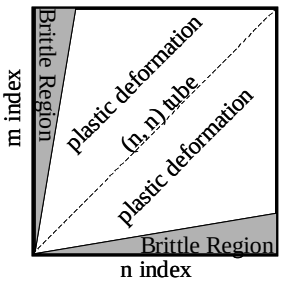


Figure 18. Ductility phase diagram for carbon nanotubes adapted from [132]. Brittle behavior is expected as zigzag geometries are approached due to the unfavorable energetics of defect generation through C–C bond rotation.

activation energies for the bond defects. However, it provides an important roadmap for investigation.

In an attempt to quantify the appearance of such defects in SWNTs under applied strain, Zhao et al. investigated the relevant activation energies using large-scale quantum calculations [133]. The generation of 5–7–7–5 defects at large applied strains was investigated for both (n, n) and $(n, 0)$ tubes—specifically, the activation barriers for defect generation computed by constrained relaxation along the simplest kinetic pathway leading to defect formation. The resulting activation energies were prohibitively large, and ranged from 8.6 eV for a (5, 5) SWNT to 9.9 eV for a (9, 0) SWNT at zero strain. Even at strains as large as 10% (far beyond the breaking strain observed in laboratory experiments), the 5–7–7–5 activation energies exceeded 4 eV, implying a vanishingly small defect generation density at or near room temperature. Consequently, it was concluded that such bond rotation in strained NTs (near room temperature) is unlikely due to the energetics of the problem. The authors note that prior agreement between estimates of the 5% strain required for 5–7–7–5 defect stability and the experimentally observed breaking strain is most likely fortuitous, and note that, for some tubes, the breaking strain exceeds 5%. A more likely scenario describing the mechanical failure of actual SWNTs and MWNTs is the nucleation of yield or failure at frozen-in defects resulting from the growth process. As the field of nanomechanics of nanotubes progresses, it will be critical to further address the roles of defects and defect generation with respect to the inelastic processes.

The validity of continuum elastic theory to treat NTs was specifically considered by Govindjee and Sackman [100]. Their primary result rested on conventional assumptions made by a number of experimental and theoretical works concerning the calculation of the moment of inertia for a tube. Through a rigorous analysis of the error associated with incorrect geometrical assumptions regarding INT, it was shown that errors in E exceeding 200% were possible. The authors suggested a geometrical correction to avoid such errors. With the exception of the geometrical subtleties in correctly calculating inertial moments, the physical validity of linear (elastic) response to describe small deformations in NTs was not challenged. Harik also investigated the applicability of classical "bending beam mechanics" to NTs based on a length scale analysis [134]. Basically, NTs were separated into two classes. For small diameter/length ratios and

for NT radii large compared to the hexagonal atomic lattice spacing, it was appropriate to apply beam-like shell theory. For long NTs with radii on the order of the atomic lattice spacing, a solid beam approach was warranted. A scaling analysis of buckling was presented to determine the dependence of critical buckling strain with the NT aspect ratio as a function of the NT diameter.

Due to the challenges in the scanning force microscopy of high aspect-ratio structures, very little direct nanomomechanical mechanical imaging has been carried out on nanotubes. Very recently, Ajayan and Geer have undertaken UFM imaging of chemical-vapor-deposited MWNTs [84] to investigate the axial variation of MWNT rigidity associated with volume defects (axial variation of tube radius). Initial results of this work are shown in Figure 19. The MWNT has a nominal diameter of 63 nm, determined by AFM height measurements. The apparent 200 nm width results from tip convolution. The axial variation of the UFM response is axially symmetric over the central uniform region of the tube. The UFM contrast lines parallel to the tube axis are due to variations in the tip–tube contact area during the scan. Near the edges of the scan, however, there are significant variations of the UFM response associated with the radius variations of the volume defects. These variations, in contrast, are associated with modulation of the mechanical response due to the variations in tube radius. It is clear, as noted by Gao et al. [121], that such volume defects could significantly alter the mechanical response (effective Young's modulus) of the tube. Investigations into quantitative modulus extraction from such measurements are continuing.

4. NANOTUBE COMPOSITES

Clearly, nanotubes (carbon or otherwise) present a class of materials of extremely high modulus and, in some cases, extremely high yield strengths. Exploitation of these nanomechanical properties has been most aggressively sought in the field of nanocomposites to extend the performance of conventional carbon fiber composites. In addition, the unique electrical and thermal conduction properties of NTs provide other avenues for nanocomposite material design. From a purely mechanics point of view, the critical parameter for composite performance is the interfacial stress transfer ability (sustainable load transfer) between the matrix and the reinforcing component (typically fibrous) [135]. This elevates materials processing issues, including NT alignment, NT dispersion, and NT/matrix interfacial adhesion to a level equal to that of NT materials properties. Early studies noted poor adhesion between NTs and an epoxy matrix [84, 136] following cutting of a crude composite with a microtome. The NTs were essentially pulled from the composite during the process, and effectively aligned. While highlighting the strength of carbon NTs, this early observation underscored material challenges in NT-based nanocomposite engineering.

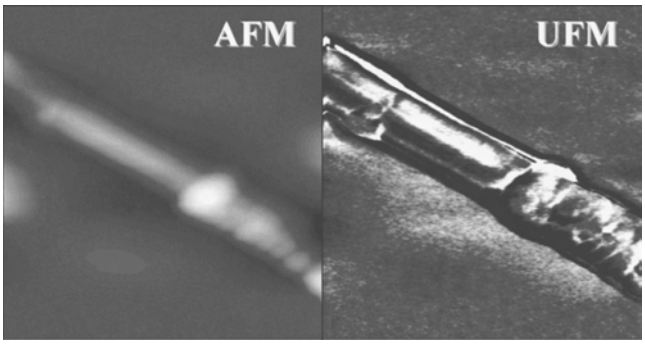


Figure 19. AFM (left) and UFM (right) image data of a carbon MWNT deposited via chemical vapor deposition (1 μm $\times$ 1 μm scan). The nominal diameter of the uniform section is 63 nm (image not corrected for tip convolution). The tube consists of a central uniform region and two defected regions near the edges of the scan. The defected regions are characterized by axial variations of the tube radius (so-called volume defects [121]). Significant modulus variations (apparent in the UFM scan) accompany these defects.

Wagner et al. investigated the nanotube–matrix stress transfer efficiency of composites consisting of arc-deposited MWNTs dispersed in a urethane/diacrylate oligomer prior to curing [137]. Their results indicated an increase in this stress transfer efficiency by over an order of magnitude compared to conventional carbon fiber-based composites. This increase is attributed to a possible "2 + 2" cycloaddition reaction occurring at the polymer/NT interface. Fragmentation of nanotubes within the composite during stress testing lends credence to the high interfacial adhesion in this system. Characterization of the load-transfer ability of carbon nanotubes was also investigated by Schadler and coworkers [138]. The optimization of load transfer requires a high interfacial shear modulus to effectively transfer the applied load to the fiber. Important factors in optimizing this modulus are micromechanical interlocking between fiber and matrix, and bonding (either covalent or van der Waals) between the two. For NTs (5% by weight) dispersed in a Shell Epon 828 epoxy resin, Schadler noted an increase in the tensile modulus E_T from 3.1 GPa for the epoxy alone to 3.78 GPa for the composite. The compression modulus exhibited a larger increase from $EC = 3.63$ to 4.5 GPa. The hysteretic nature of the stress–strain curves in the composite is attributed to either poor bonding between the matrix and the outer surface of the nanotube, or slipping within individual MWNTs, resulting in load transfer to only the outermost layer. Compared to the earlier work of Wagner et al. [137], the latter was presumed responsible.

For composites requiring high uniaxial strength, NT alignment and dispersion are critical parameters. Anisotropic mechanical stress was shown to align NTs in a thermoplastic polymer. Zhou and co-workers dispersed MWNTs produced via the arc-discharge method into a chloroform/polyhydroxyaminoether (PHAE) mixture [139]. Anisotropic stress was applied to the postdried composite mixture (50% NT by weight) at 373 K. After cooling and release, the NT orientation distribution was characterized via X-ray diffraction and transmission electron microscopy. The former revealed an NT mosaic of approximately $\pm 23^\circ$ roughly parallel to the composite stretching direction. Internal fractures in the composite displayed NT pull out similar to previous work, highlighting the need for optimization of NT/composite adhesion in the PHAE/NT system.

Composite films consisting of polyvinyl–alchohol (PVA) and MWNTs were studied by Shaffer and Windler [140].

Using short-fiber composite theory, they inferred an elastic modulus of the MWNTs of 150 MPa, approximately three orders of magnitude lower than direct measurements. This implies poor stress transfer properties of the matrix/NT interface, most likely driven by the polymer system considered. In contrast, Qian et al. [141] examined a polystyrene composite containing 1% by weight carbon MWNTs. The elastic stiffness and tensile strength of the NT composite increased 39 and 25%, respectively. Similar to other studies of NT composites, fracture in the NT/polystyrene system revealed NT pull out from the matrix. Increases in composite performance were also observed by Andrews et al. for 5% (by weight) SWNTs dispersed in pitch precursors (Ashland A500 petroleum pitch) [142]. Increases in the tensile strength from 500 MPa (pure pitch) to 850 MPa (pitch + SWNTs) were observed. The measured modulus values were 35 and 80 GPa without and with SWNTs, respectively.

As noted, stress transfer is the key figure of merit for the exploitation of nanotube mechanical properties in composites. Hence, the optimization of matrix/nanotube adhesion is critical. Liao and Li carried out theoretical investigations of a double-wall nanotube/polystyrene system using classical elasticity theory and molecular mechanics simulations to investigate this [135]. In their study, no covalent bonding was specifically considered. Rather, the simulations implied that electrostatic and van der Waals interactions between the polystyrene molecules and the nanotube surface dominated the adhesion energy. These authors also noted that mechanical deformation of the nanotube in the matrix will contribute to the adhesion. Their work pointed to mismatches in the coefficient of thermal expansion (CTE) as an important source of such deformation. Simulations of composite "cooling" from the melt implied that CTE mismatches were the primary cause for stress-induced nanotube deformations that resulted in corresponding adhesion improvements. A possible experimental route to optimize electrostatic pathways to increase NT/matrix adhesion was taken by Gong et al. [143]. That group utilized nonionic surfactants to improve the processing of NTs in an epoxy matrix with respect to dispersion and interfacial adhesion. An increase in the elastic modulus of 30% was seen; however, no comparative conclusions could be drawn regarding the modification of interfacial adhesion.

One aspect not already discussed is the effect of the multiwall nature of the tube during deformation in a nanocomposite matrix. Simplistic models assume a uniform "nonslip" bending of MWNTs in elastic composites. However, intershell slipping may significantly alter the mechanical response of the composite. This issue was considered in detail by Ru [144]. Using a multiple-shell model to study the compressed buckling of MWNTs, Ru incorporated a direct van der Waals coupling between the interior NT walls. Compared to an equivalent "single-wall" shell, the axial buckling strain for the multiple-shell model exhibited significant degradation (proportional to $1/N$, where N is the wall number). This degradation was attributed to interwall slips during bending. Although such an effect has not been demonstrated experimentally, it is an extremely important result in mapping the potential performance of MWNTs in composite applications.

Nanocomposites have also been exploited for other aspects of nanomechanical research. Li et al. used a nanocomposite to infer the elastic modulus of SWNT ropes [145]. Such ropes were dispersed in a polyvinyl chloride resin, and subjected to standard mechanical testing. Similar to other studies, nanotube pull out was a common feature of the fracture surface, and it was noted that the stress transfer from the SWNT bundles to the matrix resin was poor. From these experiments, average tensile strengths of the composites, the SWNT ropes, and the individual SWNTs were measured at 3.6, 7.5, and 11.5 GPa, respectively, on the basis of the rule of mixtures. The anomalously low value of the latter quantity is more indicative of the poor adhesion at the resin/NT interface rather than the true mechanical property of the SWNT itself.

5. NANOMECHANICS OF NANOPOROUS MATERIALS FOR DIELECTRIC APPLICATIONS

Some of the most interesting materials under investigation for nanotechnology applications are nanoporous materials. The field of nanoporous silicates is well established and far ranging. Leading applications for such materials have included nanotemplating and catalysis development. The latter has generated massive research efforts because of the exceptionally high surface-to-volume ratio of these materials due to the high pore densities achievable. For such applications, the mechanical response of these materials has been of secondary importance compared to chemical reactivity and pore morphology. Recent research in the integrated-circuit community has modified this priority. Nanoporous materials composed of silicate or polymeric materials have been engineered to possess extremely low dielectric constants [146]. For vapor-deposited silicon oxide ($\varepsilon \sim 4$), dielectric constants less than 1.8 can be achieved. This represents an incredible improvement with respect to RC dissipation in an IC interconnect line, especially for IC chips in which the interconnect line width is <130 nm. In that case, the time that it takes to establish a high-voltage pulse on a 'word" or "bit" line often exceeds the turn-on time for the associated transistor. Consequently, the capacitance of the interconnect can be the performance limiter of the entire chip [147].

Although the dielectric properties of nanoporous materials are attractive for such applications, the mechanical properties are inadequate, and represent an obstacle for integration into an IC chip. This has driven the need for nanoscale mechanical characterization to document the elastic modulus of these materials, not only for bulk values of E or E_r, but also for process-induced variations in mechanical properties. Early in the development of nanoporous silica for IC applications, Smith et al. reported moduli, as determined by isostatic compression, as low as 1 MPa for high pore density and low silicate density [148]. Compared to the 110 GPa modulus of the Cu lines to be embedded in such a dielectric matrix, the mechanical weakness of such nanomaterials presented a significant challenge. Such materials were clearly not suitable for the rigors of IC processing. As the development of nanoporous materials for this

application has matured, more typical values for porous silicate bulk Young's moduli are closer to 2 GPa, measured by nanoindentation [149–151]. From a measurement technique point of view, however, the use of nanoindentation must be considered carefully. The dramatic indentation deformation of nanoindentation for such soft materials limits the accuracy of the technique, especially for such thin films. It has been pointed out by Baklanov and co-workers that such errors may be as large as 500%, depending on the material studied [150, 152]. They reported Brilliouin light scattering and surface acoustic wave measurements of the mechanical properties of nanoporous organosilicate films as a function of porosity. The films were produced by spin coating and curing methylsilsesquioxane (MSSQ) polymers using a processing protocol designed to vary porosity from 10 to nearly 40%. The composition of the cured nanoporous films is dominated by the silicon oxide polymeric matrix, although organic moieties are still incorporated, depending on the length of cure and the connectivity of the pores. The Young's modulus measured from both Brillouin scattering and a surface acoustic wave technique were in reasonable agreement, and varied from approximately 1.8 GPa at 16% porosity to 0.8 GPa at 38% porosity. However, corresponding Young's modulus measurements carried out on the same materials using nanoindentaiton yield values that range from 3.5 to 2 GPa over the same porosity range. The dramatic error was attributed to incorrect analysis of the nanoindentation data when applied to such thin films ($\sim$0.6 μm). Specifically, these authors cite the inadequacy of current modulus extraction algorithms that were developed for stiff films on stiff substrates or stiff films on soft substrates to accurately extract moduli for thin, soft films on stiff substrates [150, 152].

More conventional mechanical measurements of porous dielectrics have also been carried out. Kloster et al. investigated the modulus, hardness, and interfacial adhesion of porous organosilicate glass thin films for potential use as low-k dielectric films for integration in IC chips as an insulator for metal interconnect lines [153]. Nanoindentation was used to acquire modulus and hardness data, and a conventional four-point bend configuration [154] was used to extract interfacial adhesion energy. The nanoindentation sampling depth was kept below 150 nm to try to avoid substrate stiffening, as noted by Baklonov. Nanoporosity of the films varied from 0 to 60%. Pore size was not reported, although it was expected to be less than 4 nm. The Young's modulus varied from 8.7 GPa for a nonporous organosilicate film (0.6 μm in thickness) to less than 0.5 GPa for a 60% porous film. Hardness values exhibited a far more precipitous decrease from 1.59 to 0.07 GPa over the same porosity range. Recall that nanoindentation hardness is characterized by the load at which the subsurface pressure at the plastic deformation below the indentor tip ceases to increase. At this point, indentation proceeds via plastic deformation over the entire deformation field [9]. Hardness values are extremely sensitive to the model used, but usually are assumed proportional to yield strength, and are measured accordingly. Consequently, for such thin soft films as those reported in [153], large uncertainties should be assumed.

Experimentally, porosity must be distinguished from pore distribution to accurately characterize the mechanical properties of nanoporous materials [155]. Recently, Aoi et al. have carried out a finite-element analysis of the mechanical properties of nanoporous materials, with emphasis on the prediction of nanomechanical properties, taking into account pore distribution [156]. For inorganic nanoporous materials (silicate based), their analysis predicted two specific regimes of linear decrease of modulus with increasing porosity (i.e., decreasing density) for random pore orientation. Specifically, for silicon oxide, a modulus of 70 GPa is expected for a maximum density of 2.5 g/cm^3 (0% porosity). For a pore size on the order of 2–4 nm, the modulus will decrease to less than 1 GPa for 50% porosity. In this regime, it is expected that little or no pore aggregation is occurring. Beyond 50% porosity, however, the decrease in slope virtually disappears as additional pore aggregates form. Hence, a significant change in the relation between porosity and modulus occur when pore distribution (i.e., aggregation) occurs. When viewed on a semi log–log plot of relative modulus as a function of relative density, their data were approximately linear, implying a power-law behavior. This behavior was summarized via the following law [156]:

$$E_{\text{por}} = E_{\text{solid}} \left\{ \frac{\rho_{\text{por}}}{\rho_{\text{solid}}} \right\}^{(3+(2\alpha/3))} \tag{23}$$

where E is the conventionally defined Young's modulus, ρ represents the density of the respective materials, and α is a pore aggregation factor. The qualitative interpretation of this law is straightforward. If nanopores are disordered, α characterizes how quickly pore aggregation will affect the modulus for a specific value of density. If pore aggregation occurs at a relatively high density, the effective modulus will be normalized downward very quickly. If, however, the nanopores are ordered such that pore aggregation is suppressed, then α vanishes, and the scaling of the modulus with density will be maintained for all film densities. This relationship between pore aggregation and density has been verified experimentally, yielding values for α around 3.

Preliminary surface nanomechanical imaging studies of porous dielectric films have been undertaken by Shekhawat

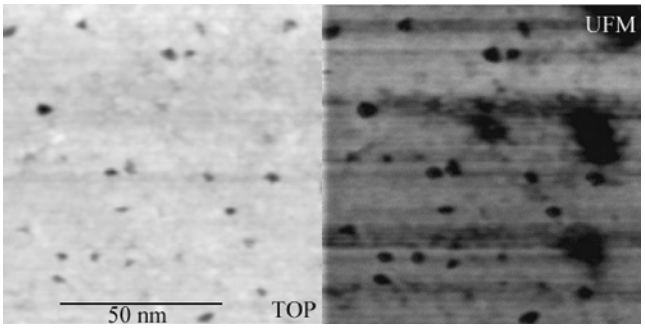


Figure 20. AFM (left) and UFM (right) image data from a nanoporous silicate film (100 nm $\times$ 100 nm scan area). Nanodepressions are evident in the AFM scan. Taking into account tip convolution effects, some of these nanodepressions are less than 2 nm in diameter. The extent of the nanodepressions is slightly increased in the UFM scan. Note the larger regions of depressed stiffness (low-contrast areas) evident in the UFM image. These features are totally absent in the AFM images, and are identified with reduced stiffness due to compositional nonuniformity or subsurface defect aggregation.

and co-workers [157]. Representative data are shown in Figure 20. The surface topographical features correspond to nanometer-sized depressions. Such features have not been previously reported, although variations in the processing of individual films with respect to cure and organic volatilization may be responsible. The UFM image corresponding to this topography is intriguing. The nanodepressions possess a reduced elastic response in the immediate vicinity. Also, a large area of reduced mechanical compliance is observed. There is no corresponding feature in the topography. This "soft area" has been tentatively identified as a region of high subsurface defect density. In an attempt to use ion bombardment to "harden" nanoporous silica films, Bakhru and co-workers investigated the effects of a 50 keV Ar ion flux on the mechanical properties of 1.0 μm thick Xerogel® films [158]. An ion dose of 10^{16} Ar ions/cm^2 was delivered to the film. Preliminary nanoindentation data suggest a hardening of the films by 30%. However, significant sample-to-sample variability was noted. To clarify the potential effects of the film nanostructure on the modification of the Xerogel nanomechanical properties, AFM and UFM imaging were undertaken. These data are shown in Figure 21. Topography measurements reveal a significant increase in surface roughness following Ar ion exposure from 0.3 nm (rms) to 3.4 nm (rms). More interestingly, however, is the effect of Ar ion bombardment on individual Xerogel grains. Comparison of the UFM images before and after ion implantation shows a dramatic increase in local compliance within certain domains and slight decreases in others. The localized modification of the Xerogel domain mechanical response may result from ion-induced bond breaking of the silicate polymer matrix, but at this time, there is no clear explanation why such effects would modify the mechanics of the local domain structure, as shown in the figure. Clearly, the quantitative application of nanomechanical imaging in materials such as these will provide a greater insight into their nanoscale properties.

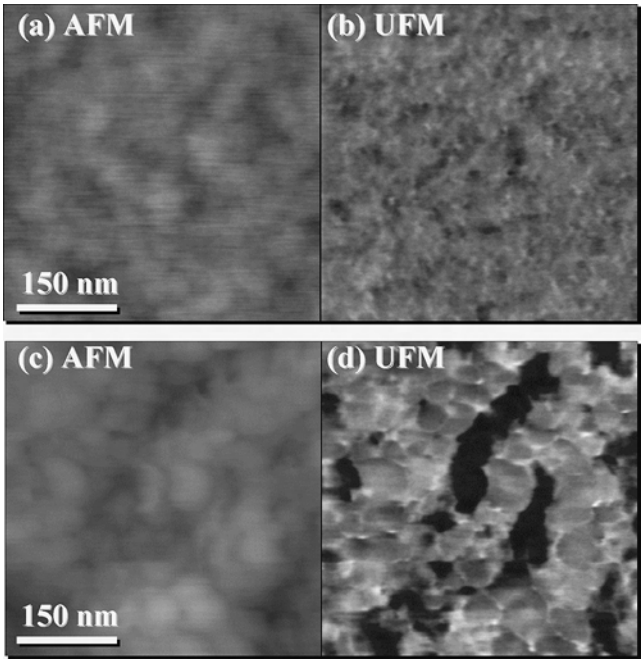


Figure 21. (a) AFM image of unimplanted Xerogel film (1 μm thickness, 0.3 nm rms roughness). (b) UFM image corresponding to the topography in (a). (c) AFM image of Xerogel film following implantation of 50 keV Ar ions (10^{16}/cm^2). The rms roughness has increased to 3.4 nm, and the domain structure has coarsened considerably. (d) UFM image of the same area reveals a significant modification of the mechanical compliance of the Xerogel domains. Note the dramatic softening of certain domains (low contrast) and stiffening of others (higher contrast).

6. CONCLUDING REMARKS

The ability to synthesize and characterize nanoscale materials and devices has opened broad avenues for nanoscience endeavor and nanotechnology development. As with the vast majority of research fields, more questions remain than have been answered concerning the fundamental degree of control of nanoscale properties and the extent to which this control can be exploited technologically. The field of nanomechanics of nanomaterials is no different. Classical approaches for describing nanomechanics have remained remarkably robust, but the limitations have already become apparent with respect to defining classical dimensions to quantum-mechanically defined objects, that is, the wall thickness of a nanotube or the true contact area of a nanotip interacting with a surface. As the scale of nanomechanical devices and imaging techniques approaches 1 nm and beyond, it might well be expected that entirely new formulations will be required.

GLOSSARY

Acoustic atomic force microscopy (AAFM) An implementation of SPM where the variation of the flexural resonance spectrum of an SPM cantilever as it is scanned over a sample surface is used to extract local mechanical properties from that surface.

Acoustic microscopy A nondestructive analytical technique whereby acoustic waves are transmitted or reflected from a material using acoustic lenses to elucidate mechanical properties, mass density, or acoustic properties of materials.

Atomic force microscopy (AFM) An implementation of SPM wherein the deflection of a cantilevered tip scanned over a sample surface is converted to a surface topography map, typically, with subnanometer out-of-plane resolution.

Bending modulus (κ) The second derivative of surface energy density with respect to curvature. For a tube, $\kappa = Ea/(1-\sigma^2)$, where σ is Poisson's ratio and a is the tube wall thickness.

Bulk modulus (B) Typically defined as the ratio of a hydrostatically applied pressure to the resulting fractional change in volume.

Cantilevered tip A pyramidal tip placed at the end of a triangular or rectangular (diving board) micromachined silicon cantilever. Used as surface probes in SPM. Typically, the radius of the tip is on the order of 10 nm. Spring constants of cantilevered tips range from 0.01 to 10 N/m.

Contact stiffness (S) The first derivative of a surface-restoring force with respect to out-of-plane surface displacement.

Continuum elastic theory Fundamental theory of the deformation of solids. Assumes that a system can be approximated as a completely continuous body, and that deformations due to external stresses are elastic (completely reversible).

Crystal defects Departures from a perfect, periodic structure in a crystal lattice. Typical defects include dislocations (extra or missing rows or planes of atoms), vacancies (missing atoms), interstitials (additional atoms between lattice sites), or disclinations (coordination number defects).

Density functional theory (DFT) An analytic or numerical modeling technique (often for an electronic structure) wherein a local electronic or probability function is variationally optimized to calculate the ground or excited states of a system.

Derjaguin, Muller, and Toporov theory (DMT theory) Theory of nanoscale deformation of a tip at an adhesive surface wherein surface adhesive forces are restricted to a fraction of the contact area.

Elasticity The quality of elastic response: linear dependence of restoring force on displacement.

Finite-element analysis (FEA) A numerical modeling technique for which a continuous system (solid or fluid) is represented by a 3-D mesh or matrix of finite-sized elements. The response of the entire system to external stresses, temperature changes, applied fields, and so on is determined by solving classical equations of motion (Newton's equation, Maxwell's equations, Navier–Stokes equation, etc.) at the interfaces of each element.

Force modulation microscopy (FMM) An implementation of SPM wherein the oscillation amplitude of a cantilevered tip in contact with a substrate is measured as the tip is scanned over a surface. For surfaces exhibiting a contact stiffness below the cantilever spring constant, such scans can reveal the spatial variation of surface mechanical properties.

Graphene Referring to a hexagonal atomic structure of carbon as in a single a–b plane of graphite.

Hertzian contact Model of a surface–surface contact for which only contact forces (short-range electrostatic forces) are considered. Long-range and adhesive forces are ignored.

Heterodyne force microscopy (HFM) An implementation of SPM wherein heterodyne coupling of sample and cantilever ultrasonic vibrations (via the nonlinear-force displacement relationship of a scanning probe tip at a surface) is used to probe the nanometer spatial variations of surface visoelastic properties.

Integrated circuits (ICs) Semiconducting devices and associated electrical interconnects integrated on a single substrate via multilayer, photolithography-based processing.

Johnson–Kendall–Roberts–Sperling theory (JKRS theory) Theory of nanoscale deformation of a tip at an adhesive surface assuming the validity of classical elastic theory.

Local density of states (LODS) In the context of the electronic properties of a material, the LDOS is the number of quantum mechanical states per unit volume that can be occupied by an electron.

Molecular dynamics (MD) Numerical modeling paradigm wherein an empirical or analytically derived potential is used to calculate atomic or molecular configurations based on the numerical solutions to the Newtonian or quantum mechanical equations of motion.

Mechanical resonance Dynamic condition wherein a mechanical system is driven at a natural vibration frequency.

Microelectrical mechanical systems (MEMS) Term typically applied to 2-D or 3-D microstructures fabricated using the techniques of integrated-circuit fabrication.

Nanocomposite A multicomponent composite material where the spatial size or modulation of at least one of the components is characterized by a nanometer-length scale.

Nanoindentation Mechanical characterization technique wherein material modulus, hardness, and fracture strength are measured by indentation with a pyramidal or tetrahedral tip. The indentation is typically on the order of microns.

Nanomaterials Engineered materials exhibiting unique properties resulting from the manipulation of the material structure or composition at the nanometer-length scale.

Nanomechanical imaging Direct imaging of mechanical properties (modulus, visoelasticity, etc.) with nanometer-scale resolution via near-field microscopy.

Nanomechanics The mechanics of nanometer-scale contacts and the deformation of nanoscale structures.

Nanorope Term typically applied to bundles of nanotubes in a hexagonal arrangement forming a rope-like braid.

Nanotribology The study of tribology (frictional interactions of surfaces) at the nanometer-length scale.

Nanotube A nanometer-scale tube formed by wrapping single or multiple atomic layers in a cylindrical geometry. The former is referred to as a single-wall nanotube (SWNT). The latter is referred to as a multiwall nanotube (MWNT). Diameters range typically from 2 to >100 nm. Typically formed by graphene sheets of carbon nanotubes. Can also be formed by a number of other inorganic compounds.

Organosilicates A term typically applied to solids or chemical compounds involving siloxane groups with organic ligands.

Piezoelectric The property of a material lacking inversion symmetry whereby application of physical stress results in a change in the polarization of the material. Conversely, application of an external voltage to a piezoelectric material will induce a spatial deformation.

Plasticity Nonelastic response of a material to an applied stress wherein a nonrecoverable (plastic) deformation is induced.

Poisson's ratio (σ) Ratio of the transverse compression to the longitudinal extension of an object subjected to a tensile stress.

Quantum dot (QD) Term typically referring to nanometer-scale semiconductor particle (synthesized chemically in nanopowder form or grown on an appropriate substrate) exploited for unique electrical or optoelectronic properties due to quantum-mechanical localization phenomena.

Reduced Young's modulus ($E_r \equiv E/(1-\sigma^2)$) Conventional Young's modulus normalized by Poisson's ratio.

Shear modulus (μ) Typically defined as the ratio of the applied shear stress to the fractional transverse displacement of an object.

Scanning local acceleration microscopy (SLAM) An implementation of SPM whereby the deflection of the cantilevered tip is proportional to the tip acceleration, and can be related to the mechanical properties of the substrate.

Scanning probe microscopy (SPM) Any number of techniques in which a local probe is rastered across a surface with subnanometer precision to characterize topographic or material properties. Examples include scanning tunneling microscopy (STM), in which local tunneling currents between a scanning tip and a substrate are used to map the local density of electronic states, or atomic force microscopy (AFM), where the local deflection of a cantilevered tip by contact or van der Waals forces is recorded as the tip is scanned over a surface to measure surface topography.

Tensile strength Typically defined as the longitudinal force or stress applied to a material for which nonreversible deformation results.

Tight-binding approximation (TB) An approximation for solving quantum-mechanical equations of motion for a material in which it is assumed that electrons are localized near the parent ion.

Ultrasonic force microscopy (UFM) An implementation of SPM wherein an out-of-plane ultrasonic vibration is applied to a sample on which a cantilevered tip (in contact) is scanned. For sufficient vibration amplitude, a tip-deflection signal proportional to the local sample modulus is obtained.

van der Waals forces Weak, long-range forces resulting from induced electric dipole fields. Typically varies inversely as the separation to the sixth power.

Viscoelasticity Dynamics of elastic deformation, including the temporal response of materials to an elastic deformation.

Xerogel Trade name referring to a nanoporous silicate material.

Young's modulus (E) Typically defined as the ratio of the applied longitudinal stress to the fractional I longitudinal deformation of an object.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge the contributions and support of many colleagues, students, and sponsors in assembling this review.

REFERENCES

1. "National Nanotechnology Initiative: Leading to the Next Industrial." Interagency Working Group on Nanoscience, Engineering and Technology, White House Press, Washington, DC, 1999.
2. G. Frazier, "An Ideology for Nanoelectronics in Computation." Plenum, New York, 1988.
3. S. Iijima, *Nature* 354, 56 (1991).
4. L. D. Landau and E. M. Lifshitz, "Theory of Elasticity." Butterworth Heineman, Oxford, 1998.
5. Z. L. Wang, P. Poncharal, and W. A. de Heer, *J. Phys. Chem. Solids* 61, 1025 (2000).
6. Q. Zhao, M. Buongiorno Nardelli, and J. Bernholc, *Phys. Rev. B* 65, 144105 (2002).
7. G. A. D. Briggs, "Acoustic Microscopy." Clarendon, Oxford, 1992.
8. E. Meyer, R. M. Overney, K. Dransfeld, and T. Gyalog, "Nanoscience: Friction and Rheology on the Nanometer Scale." World Scientific, Singapore, 1998.
9. A. C. Fischer-Cripps, "Nanoindentation." Springer, New York, 2002.
10. H. Hertz, "Hertz's Miscellaneous Reports." Macmillan, London, 1896.
11. H. Hertz, *Verh. Ver. Beforderung Gewerbe Fleisses* 61, 410 (1882).
12. F. J. Giessibl, *Phys. Rev. B* 56, 16010 (1997).
13. N. Sasaki and M. Tsukada, *Appl. Surf. Sci.* 140, 339 (1999).
14. M. Tsukada, N. Sasaki, R. Tamura, N. Sato, and K. Abe, *Surf. Sci.* 401, 355 (1998).
15. J. P. Spatz, S. Sheiko, M. Moller, R. G. Winkler, P. Reineker, and O. Marti, *Nanotechnol.* 6, 4 (1995).
16. J. Chen, R. K. Workman, D. Sarid, and R. Hoper, *Nanotechnol*. 5, 199 (1994).
17. K. L. Johnson, K. Kendall, and A. D. Roberts, *Proc. R. Soc. Lond. A* 324, 310 (1971).
18. G. Sperling, Eine Theorie der haftung von Festsoffteilchen an festern Korpern Ph.D. Dissertation, Fakultat der Maschinenwesen T. H. Karlsruhe, 1964.
19. K. Inagaki, O. Matsuda, and O. B. Wright, *Appl. Phys. Lett.* 80, 2386 (2002).
20. B. V. Derjaguin, V. M. Muller, and Y. P. Toporov, *J. Coll. Interface Sci.* 53, 314 (1975).
21. G. Binning, H. Rohrer, Ch. Gerber, and E. Weibel, *Phys. Rev. Lett.* 49, 57 (1982).
22. A. Baratoff, G. Binnig, and H. Rohrer, *J. Vac. Sci. Technol. B* 1, 703 (1983).
23. N. Garcia, C. Ocal, and F. Flores, *Phys. Rev. Lett.* 50, 2002 (1983).
24. G. Binnig, C. F. Quate, and Ch. Gerber, *Phys. Rev. Lett.* 56, 930 (1986).
25. H. Geisler, M. Hoehn, M. Rambach, M. A. Meyer, E. Zschech, M. Mertig, A. Romanov, M. Bobeth, W. Pompe, and R. E. Geer (unpublished).
26. N. A. Burnham and R. J. Colton, *J. Vac. Sci. Technol. A* 7, 2906 (1989).
27. P. Maivald et al., *Nanotechnol*. 2, 103 (1991).
28. S. P. Jarvis and J. B. Pethica, in "Forces in Scanning Probe Methods." Kluwer Academic NATO ASI Series 286, Kluwer Academic, Dordrecht, 1995.
29. E.-L. Florin, M. Radmacher, B. Fleck, and H. E. Gaub, *Rev. Sci. Instrum.* 65, 639 (1994).
30. B. Cretin and F. Sthal, *Appl. Phys. Lett.* 62, 829 (1993).
31. N. A. Burnham, A. J. Kulik, G. Gremaud, P.-J. Gallo, and F. Oulevey, *J. Vac. Sci. Technol. B* 14, 794 (1996).
32. N. A. Burnham, O. P. Behrend, F. Oulevey, G. Gremaud, P.-J. Gallo, D. Gourdon, E. Dupas, A. J. Kulik, M. M. Pollock, and G. A. D. Briggs, *Nanotechnol*. 8, 67 (1997).
33. O. Kolosov and K. Yamanaka, *Jpn. J. Appl. Phys.* 32, L1095 (1993).
34. K. Yamanaka, H. Ogiso, and O. Kolosov, *Appl. Phys. Lett.* 64, 178 (1994).
35. F. Dinelli, S. K. Biswas, G. A. D. Briggs, and O. V. Kolosov, *Phys. Rev. B* 61, 13995 (2000).
36. Y. Zheng (unpublished results).
37. N. A. Burnham, A. J. Kulik, G. Gremaud, and G. A. D. Briggs, *Phys. Rev. Lett.* 74, 5092 (1995).
38. G. S. Shekhawat, O. V. Kolosov, G. A. D. Briggs, and R. E. Geer, *J. Appl. Phys.* 91, 4549 (2002).
39. F. Dinelli, M. R. Castell, D. A. Ritchies, N. J. Mason, G. A. D. Briggs, and O. V. Kolosov, *Philos. Mag. A* 80, 2299 (2000).

40. L. Muthuswami, E. S. Moyer, Z. Li, and R. E. Geer, "Proceedings of the IEEE 2002 International Interconnect Technology Conference," Cat. No. 02EX519C, 2002, p. 239.
41. K. Yamanaka, H. Ogiso, and O. Kolosov, *Jpn. J. Appl. Phys.* 33, 3197 (1994).
42. G. G. Yaralioglu, F. L. Degertekin, K. B. Crozier, and C. F. Quate, *J. Appl. Phys.* 87, 7491 (2000).
43. H. Geisler, M. Hoehn, M. Rambach, M. A. Meyer, E. Zschech, M. Mertig, A. Romanov, M. Bobeth, W. Pompe, and R. E. Geer, "Proceedings of the XIIth International Conference on Microscopy of Semiconducting Materials," 2001.
44. B. Altemus, G. Shekhawat, X. Bai, R. E. Geer, and J. Castracane, *Proc. SPIE, Int. Soc. Opt. Eng*. 4558, 143 (2001).
45. O. V. Kolosov, M. R. Castell, C. D. Marsh, G. A. D. Briggs, T. I. Kamins, and R. S. Williams, *Phys. Rev. Lett.* 81, 1046 (1998).
46. T. I. Kaimins, *J. Appl. Phys.* 81, 211 (1997).
47. G. S. Shekhwat and R. E. Geer (unpublished).
48. D. J. Eaglesham and R. Hull, *Mater. Sci. Eng. B* 30, 197 (1995).
49. M. R. Castell, D. D. Perovic, and H. Lafontaine, *Ultramicroscopy* 69, 279 (1997).
50. "Handbook of Chemistry and Physics," 66th ed. CRC Press, Cleveland, OH, 1986.
51. T. R. Stoner, University at Albany (private communication).
52. T. Stark and A. E. Kaloyeros, University at Albany (unpublished results).
53. G. Shekhawat and R. E. Geer, *Nanotechnol*. (in press).
54. G. S. Shekhawat, O. V. Kolosov, G. A. D. Briggs, E. O. Schaffer, S. Martin, and R. E. Geer, "Materials, Technology and Reliability for Advanced Interconnects and Low-k Dielectrics. Symposium: Materials Research Society Symposium Proceedings," Vol. 612, 2001, p. 1.
55. G. S. Shekhawat, O. V. Kolosov, G. A. D. Briggs, E. O. Shaffer, S. J. Martin, and R. E. Geer, "Proceedings of the International Interconnect Technology Conference," IEEE Cat. No. 00EX407, 2000, p. 96.
56. G. S. Shekhawat, G. A. D. Briggs, O. V. Kolosov, and R. E. Geer, "Characterization and Metrology for ULSI Technology 2000, AIP Conference Proceedings," Vol. 550, 2001, p. 449.
57. R. E. Geer and G. S. Shekhawat, "Proceedings of the 9th International Conference on Composites Engineering," 2002, p. 299.
58. P. Van Zant, "Microchip Fabrication: A Practical Guide to Semiconductor Manufacturing." McGraw-Hill, New York, 1997.
59. R. Talevi, H. Gundlach, Z. Bian, A. Knorr, M. van Gestel, S. Padiyar, A. E. Kaloyeros, R. E. Geer, E. O. Shaffer, and S. Martin, *J. Vac. Sci. Technol. B* 18, 252 (2000).
60. H. Gundlach, R. Talevi, Z. Bian, G. Nuesca, S. Sankaran, K. Kumar, A. E. Kaloyeros, R. E. Geer, J. Liu, J. Hummel, E. O. Shaffer, and S. J. Martin, *J. Vac. Sci. Technol. B* 18, 2463 (2000).
61. M. R. Baklonov, M. Van Hove, G. Mannert, S. Vanhaelemeersch, H. Bender, T. Conard, and K. Maex, *J. Vac. Sci. Technol. B* 18, 1281 (2000).
62. S. A. Vitale, H. Chae, and H. H. Sawin, *J. Vac. Sci. Technol. A* 18, 2770 (2000).
63. R. E. Geer, O. V. Kolosov, and G. A. D. Briggs (unpublished).
64. W. W. Graessly, "Physical Properties of Polymers," 2nd ed. ACS, Washington, DC, 1992.
65. F. Iwata, T. Matsumoto, and A. Sasaki, *Nanotechnol*. 11, 10 (2000).
66. E. Hamada and R. Kaneko, *J. Phys. D: Appl. Phys.* 25, 53 (1992).
67. K. Inagaki, O. V. Kolosov, G. A. D. Briggs, and O. B. Wright, *Appl. Phys. Lett.* 76, 1836 (2000).
68. U. Rabe and W. Arnold, *Appl. Phys. Lett.* 64, 21 (1994).
69. U. Rabe, V. Scherer, S. Hirsekorn, and W. Arnold, *J. Vac. Sci. Technol. B* 15, 1997.
70. U. Rabe, S. Amelio, E. Kester, V. Schere, S. Hirsekorn, and W. Arnold, *Ultrason*. 38, 430 (2000).
71. K. B. Crozier, G. G. Yaralioglu, F. L. Defertekin, J. D. Adams, S. C. Minne, and C. F. Quate, *Appl. Phys. Lett.* 76, 1950 (2000).
72. O. Wright and N. Nishiguchi, *Appl. Phys. Lett.* 71, 626 (1997).
73. G. G. Yaralioglu, F. L. Degertekin, K. B. Crozier, and C. F. Quate, *J. Appl. Phys.* 87, 7491 (2000).
74. T. Tsuji, H. Irihama, and K. Yamanaka, *Jpn. J. Appl. Phys.* 41, 832 (2002).
75. S. P. Jarvis, A. Oral, T. P. Weihs, and J. B. Pethica, *Rev. Sci. Instrum.* 64, 3515 (1993).
76. G. Behme and T. Hesjedal, *Appl. Phys. A* 70, 361 (2000).
77. G. Behme and T. Hesjedal, *J. Appl. Phys.* 89, 4850 (2001).
78. H. Safar, R. N. Kleiman, B. P. Barber, P. L. Gammel, J. Pastalan, H. Huggins, L. Fetter, and R. Miller, *Appl. Phys. Lett.* 77, 136 (2000).
79. K. Yamanaka, Y. Maruyama, T. Tsuji, and K. Nakamoto, *Appl. Phys. Lett.* 78, 1939 (2001).
80. O. V. Kolosov and G. A. D. Briggs, International Patent WO 98/08046.
81. M. T. Cuberes, H. E. Assender, G. A. D. Briggs, and O. V. Kolosov, *J. Phys. D: Appl. Phys.* 33, 2347 (2000).
82. G. S. Shekhawat, H. Xie, Y. Zheng, and R. E. Geer, *Mater. Res. Symp. Proc.* 2002.
83. D. H. Robertson, D. W. Brenner, and J. W. Mintmire, *Phys. Rev. B* 45, 12592 (1992).
84. P. M. Ajayan and T. W. Ebbesen, *Rep. Prog. Phys.* 60, 1025 (1997).
85. A. Thess, R. Lee, P. Nikolaev, H. Dai, P. Petit, J. Robert, C. Xu, Y. H. Lee, S. G. Kim, A. G. Rinzler, D. T. Colbert, G. E. Scuseria, D. Tomanek, J. E. Fisher, and R. E. Smalley, *Science* 273, 483 (1996).
86. C. Journet, W. Maser, P. Bernier, A. Loiseau, P. Deniard, S. Lefrant, R. Lee, and J. Fischer, *Nature* 388, 756 (1997).
87. M. F. Yu, B. S. Files, S. Arepalli, and R. S. Ruoff, *Phys. Rev. Lett.* 84, 5552 (2000).
88. L. Vaccarini, C. Goze, L. Henrard, E. Hernandez, P. Bernier, and A. Rubio, *Carbon* 38, 1681 (2000).
89. D. Tekleab and D. L. Carroll, *Phys. Rev. B* 64, 35419 (2001).
90. A. Maiti, *Chem. Phys. Lett.* 331, 21 (2000).
91. J. Bernholc, C. Brabec, M. Buongiornio Nardelli, A. Maiti, C. Roland, and B. I. Yakobson, *Appl. Phys. A* 67, 39 (1998).
92. J. P. Salvetat, J. M. Bonard, N. H. Thomson, A. J. Kulik, L. Forro, W. Benoit, and L. Zuppiroli, *Appl. Phys. A* 69, 255 (1999).
93. G. Zhou, W. Duan, and B. Gu, *Chem. Phys. Lett.* 333, 344 (2001).
94. J. W. Mintmire, B. I. Dunlap, and C. T. White, *Phys. Rev. Lett.* 68, 631 (1991).
95. L. Meirovitch, "Elements of Vibration Analysis." McGraw-Hill, New York, 1986.
96. B. I. Yakobson, C. J. Brabec, and J. Bernholc, *Phys. Rev. Lett.* 76, 2511 (1996).
97. J. Tersoff, *Phys. Rev. B* 37, 6991 (1988).
98. D. W. Brenner, *Phys. Rev. B* 42, 9458 (1990).
99. E. Hernandez, C. Goze, P. Bernier, and A. Rubio, *Appl. Phys. A* 68, 287 (1999).
100. S. Govindjee and J. L. Sackman, *Solid State Commun.* 110, 227 (1999).
101. J. P. Lu, *Phys. Rev. Lett.* 79, 1297 (1997).
102. N. Yao and V. Lordi, *J. Appl. Phys.* 84, 1939 (1998).
103. A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard, III, and W. M. Skiff, *J. Am. Chem. Soc.* 114, 10024 (1992).
104. E. J. Seldin and C. W. Nezbeda, *J. Appl. Phys.* 41, 3389 (1970).
105. A. K. Rappé, K. S. Colwell, and C. J. Casewit, *Inorg. Chem.* 32, 3438 (1993).
106. C. J. Casewit, K. S. Colwell, and A. K. Rappé, *J. Am. Chem. Soc.* 114, 10035 (1992).
107. G. Gao, T. Gagin, and W. A. Goddard, *Nanotechnol*. 9, 184 (1998).
108. D. Porezag, T. Frauenheim, T. Köhler, G. Seifert, and R. Kashner, *Phys. Rev. B* 51, 12947 (1995); J. Widany, T. Frauenheim, T. Köhler, M. Sternberg, D. Porezag, G. Jungnickel, and G. Seifert, *Phys. Rev. B* 53, 4443 (1996).
109. C. M. Goringe, D. R. Bowler, and E. Hernandez, *Rep. Prog. Phys.* 60, 1447 (1997).

110. Z. Xin, Z. Jianjun, and O.-Y. Zhong-can, *Phys. Rev. B* 62, 13692 (2000).
111. M. Menon, E. Richter, and K. R. Subbaswamy, *J. Chem. Phys.* 104, 5875 (1996).
112. J. Yu, P. K. Kalia, and P. Vashishta, *J. Chem. Phys.* 103, 6697 (1995); D. Tomanek, *Phys. Rev. Lett.* 67, 2331 (1991).
113. Z.-C. Ou-Yang, Z. B. Su, and C.-L. Wang, *Phys. Rev. Lett.* 78, 4055 (1997).
114. J. Guenzburger and D. E. Ellis, *Phys. Rev. B* 45, 285 (1992).
115. D. E. Ellis, J. Guo, and H. P. Cheng, *J. Phys. Chem.* 92, 3024 (1988).
116. M. M. Treacy, T. W. Ebbesen, and J. M. Gibson, *Nature* 381, 678 (1996).
117. A. Krishnan, E. Dujardin, T. W. Ebbesen, P. N. Yianilos, and M. M. J. Treacy, *Phys. Rev. B* 58, 14013 (1998).
118. J. P. Salvetat, G. A. D. Briggs, J. M. Bonard, R. R. Bacsa, A. J. Kulik, T. Stockli, N. A. Burnham, and L. Forro, *Phys. Rev. Lett.* 82, 944 (1999).
119. M. F. Yu, B. S. Files, S. Arepalli, and R. S. Ruoff, *Phys. Rev. Lett.* 84, 5552 (2000).
120. Z. W. Pan, S. S. Xie, L. Lu, B. H. Chang, L. F. Sun, W. Y. Zhou, G. Wang, and D. L. Zhang, *Appl. Phys. Lett.* 74, 3152 (1999).
121. R. Gao, Z. L. Wang, Z. Bai, W. A. de Heer, L. Dai, and M. Gao, *Phys. Rev. Lett.* 85, 622 (2000).
122. D. Srivastava, M. Menon, and K. Cho, *Phys. Rev. B* 63, 195413 (2001).
123. N. G. Chopra and A. Zettl, *Solid State Commun.* 105, 297 (1998).
124. R. S. Ruoff and D. C. Lorents, *Carbon* 33, 925 (1995).
125. S. Xie, W. Li, Z. Pan, B. Chang, and L. Sun, *J. Phys. Chem. Solids* 61, 1153 (2000).
126. M. F. Yu, M. J. Dyer, and R. S. Ruoff, *J. Appl. Phys.* 89, 4554 (2001).
127. M. R. Falvo, G. J. Clary, R. M. Tayulor, V. Chi, F. P. Brooks, S. Washburn, and R. Superfine, *Nature* 389, 582 (1997).
128. E. L. Axelrad, *Int. J. Non-Linear Mech.* 20, 249 (1985).
129. N. G. Chopra, *Nature* 377, 135 (1995).
130. S. Iijima, C. Brabec, A. Maiti, and J. Bernholc, *J. Chem. Phys.* 104, 2089 (1996).
131. M. Buonigiorno Nardelli, B. I. Yakobson, and J. Bernholc, *Phys. Rev. B* 57, 4277 (1998).
132. M. Buonigiorno Nardelli, B. I. Yakobson, and J. Bernholc, *Phys. Rev. Lett.* 81, 4656 (1998).
133. Q. Zhao, M. Buongiorno Nardelli, and J. Bernholc, *Phys. Rev. B* 65, 144105 (2002).
134. V. M. Harik, *Solid State Commun.* 120, 331 (2001).
135. K. Liao and S. Li, *Appl. Phys. Lett.* 79, 4225 (2001).
136. P. M. Ajayan, O. Stephan, C. Colliex, and D. Trauth, *Science* 265, 1212 (1994).
137. H. D. Wagner, O. Lourie, Y. Feldman, and R. Tenne, *Appl. Phys. Lett.* 72, 12 (1998).
138. L. S. Schadler, S. C. Giannaris, and P. M. Ajayan, *Appl. Phys. Lett.* 73, 3842 (1998).
139. L. Jin, C. Bower, and O. Zhou, *Appl. Phys. Lett.* 73, 1197 (1998).
140. M. S. P. Shaffer and A. H. Windler, *Adv. Mater.* 11, 937 (1999).
141. D. Qian, E. C. Dickey, R. Andrews, and T. Rantell, *Appl. Phys. Lett.* 76, 2868 (2000).
142. R. Andrews, D. Jacques, A. M. Rao, T. Rantell, F. Derbyshire, Y. Chen, J. Chen, and R. C. Haddon, *Appl. Phys. Lett.* 75, 1329 (1999).
143. X. Gong, J. Liu, S. Baskaran, R. D. Voise, and J. S. Young, *Chem. Mater.* 12, 1049 (2000).
144. C. Q. Ru, *J. Appl. Phys.* 89, 3426 (2001).
145. F. Li, H. M. Cheng, S. Bai, G. Su, and M. S. Dresselhaus, *Appl. Phys. Lett.* 77, 3161 (2000).
146. Y. Xu, Y. Tsai, D. W. Zheng, K. N. Tu, Chung W. Ong, C. L. Choy, B. Zhao, Q.-Z. Liu, and M. Brongo, *J. Appl. Phys.* 88, 5744 (2000).
147. "The International Technology Roadmap for Semiconductors," 1999 ed. Semiconductor Industry Association, Santa Clara, CA, 1999.
148. D. M. Smith, J. Anderson, C. C. Cho, G. P. Hohnston, and S. P. Jeng, *Mater. Res. Soc. Symp. Proc.* 381, 261 (1995).
149. C. Jin and J. Wetzel, "Proceedings of the International Interconnect Technology Conference," IEEE Cat. No. 00EX407, 2000, p. 99.
150. C. M. Flannery and M. R. Baklonov, "Proceedings of the International Interconnect Technology Conference," IEEE Cat. No. 02EX519C, 2002, p. 233.
151. S. Yang, J. C.-H. Pai, C.-S. Pai, G. Dabbagh, O. Nalamasu, E. Reichmanis, J. Seputro, and Y. S. Obeng, *J. Vac. Sci. Technol. B* 19, 2155 (2001).
152. C. M. Flannery, T. Wittkowski, K. Jung, B. Hillebrands, and M. R. Baklanov, *Appl. Phys. Lett.* 80, 4594 (2002).
153. G. Kloster, T. Scherban, G. Xu, J. Blaine, B. Sun, and Y. Zhou, "Proceedings of the International Interconnect Technology Conference," IEEE Cat. No. 02EX519C, 2002, p. 242.
154. Q. Ma, *J. Mater. Res.* 12, 840 (1997).
155. D. W. Gidley, W. E. Frieze, T. L. Dull, A. F. Yee, E. T. Ryan, and H.-M. Ho, *Phys. Rev. B* 60, R5157 (1999).
156. N. Aoi, T. Fukuda, and H. Yanazawa, "Proceedings of the International Interconnect Technology Conference," IEEE Cat. No. 02EX519C, 2002, p. 72.
157. G. Shekhawat, J. Wetzel, J. Lin, and R. E. Geer (unpublished).
158. N. A. Roy, H. Bakhru, G. S. Shekhawat, and R. E. Geer (unpublished).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanomembranes

Cees J. M. van Rijn, Wietze Nijdam

University of Twente, Enschede, The Netherlands

CONTENTS

1. MEMBRANE TECHNOLOGY

1.1. Introduction

Membranes play an essential role in nature as well as in our industrial society today. Biological membranes have an essential exchange function in the metabolism of all living species. By these exchange processes a single cell can be supplied with nutrients and the cell can excrete waste products through its outer cell membrane. Moreover, these exchange processes enable communications between different cells, by exchange of, for example, hormones, which made the development of complex life forms possible.

Considering that the first membrane experiments have already been carried out in the 18th century, using membranes (from Latin *membrana*, a thin sheet) of biological origin [1], industrial membrane separation with synthetic membranes has been strongly developed since the introduction of the asymmetric polymeric membrane [2] in the early 1960s. See Figure 1.

Most of the synthetic membranes today obtained with a phase inversion process have an inner spongelike structure which contributes to an additional fluid resistance and *a fortiori* a diminished operational product flux at a given transmembrane pressure. Also the retention of particles, proteins, and other coagulates is not only determined by the pores present in the skin layer but also by the sponge support structure itself. Moreover, in certain cases, like clarification of beverages, cleaning agents must be removed completely after cleaning to prevent contamination of the beverage. Various attempts have been made to downscale sieve-shaped filters that do not exhibit these problems, like track-etched membranes and anopore membranes (see Figs. 14 and 15 in Section 1.3). However, both membrane types lack the possibility of adjusting the pore size freely in conjunction with highest achievable porosity.

Using nano- and microtechniques originating from semiconductor technology it is possible to fabricate thin membranes with perfectly uniform pores [3]. The membrane, the microsieve, was made using flat substrates, thin-layer deposition techniques, photolithography, and high-resolution etching methods. Microsieves and nanosieves (with pore size <100 nm) may be used for many filtration applications, but they can also be used as a scaffold to study particles, microorganisms, and lipid bilayer that are retained or adhere to the surface of the micro/nanosieve. See Figure 2.

The screening of facilitated transport of a wide variety of molecules through natural and artificially constructed porins (e.g., ionchannels) in lipid cell membrane layers is an increasing object of interest. A prerequisite therefore is the construction of adequate micro- and nanobiotools obtained with novel nano- and microengineering techniques.

Current biotechnological separations (e.g., food, beverage, and pharmaceutical industry) and also membrane bioreactors are fast developing applications of artificial membranes. Also there are emerging applications, such as affinity membranes (e.g., protein purification), biosensors (e.g., glucose sensors), biohybrid organs (e.g., artificial livers), and artificial tissue structures.

Application of novel membrane structures also requires development of appropriate models for properties and processes applying nanotechnology. Properties related to design and analysis of separation processes in particular, such as adsorption, diffusion in external force fields, and other fluid parameters, play an important role here. For a good understanding of nanomembrane technology, good models for all underlying phenomena must be present.

1.2. Membrane Processes

In the chemical process industry one often encounters the problem of separating a mixture in its components. Using membranes one can in principle carry out the majority of these separation processes and may complement or form

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (47–82)

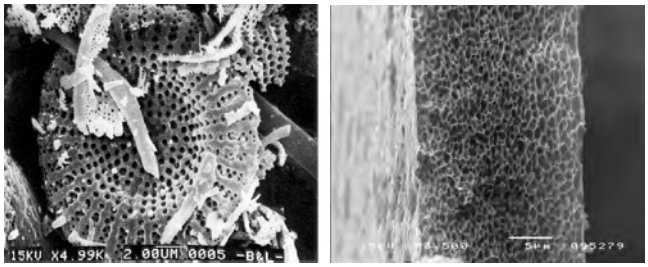


Figure 1. Scanning electron microscope (SEM) photographs. Left: a microporous skeleton of a diatomaceous earth fossil. Diatom cells are enclosed by cell walls made largely of silica (SiO_2). This glassy shell, or frustule, consists of two tightly fitting halves often resembling a flat, round, or elongated box. The minute perforations allow dissolved gases and nutrients to enter and exit. Right: cross-section of a microporous synthetic polyethersulfon membrane obtained with a phase inversion process. The membrane may have at least one outer skin layer with a mean pore size smaller than that of the inner part. Courtesy of Aquamarijn Research/Membrane Technology Group, University of Twente.

an alternative for chemical processes like distillation, extraction, fractionation, adsorption, etc. Advantages of membrane filtration are, among others, that there is a low energy consumption, the separation can be carried out continuously, up-scaling is relatively simple, and membrane technology can be used for almost any kind of separation.

The success of a membrane application or process is closely related to the intrinsic properties of the membrane. Interfacial interactions between membrane surface, surrounding environment, and solutes govern membrane performance to a great extent. These interactions have considerable impact on transport characteristics, selectivity, fouling propensity, and bio- and hemocompatibility of the membrane, especially in biotechnological and medical applications, where highly adsorptive solutes such as proteins are present and the adsorption and consecutive fouling of the membrane lead to considerable losses in flux, selectivity, and performance [5]. See Tables 1 and 2.

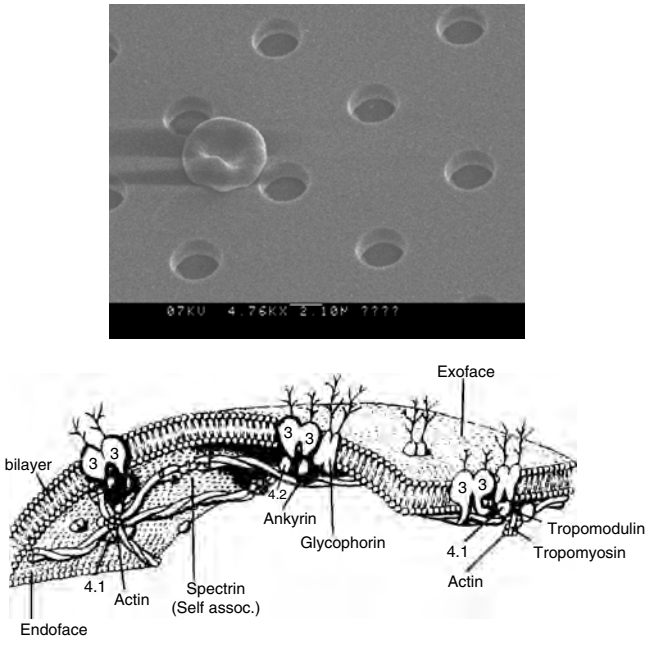


Figure 2. SEM photograph of a human blood cell retained on a microsieve membrane. Courtesy of Aquamarijn Research. Right: Drawing of the lipid bilayer (cell membrane) of a human red blood cell, showing the single-pass transmembrane protein glycophorin. Reprinted with permission from [4], S. Chien, *Microvascular Res.* 44, 243 (1993). © 1993, Academic Press.

1.2.1. Filtration through Size Exclusion

Molecular Sieves In 1756 Cronstedt discovered a selection of natural porous minerals which, when heated, produced steam. Due to this phenomenon he called them zeolites from the Greek for boiling stone (zeo from zein—to boil; lithos—stone). In 1949, Milton working at Union Carbide produced the world's first manmade zeolites, the most important being Linde A and X. Linde A has gone on to become one the most widely used zeolites. Zeolites are highly crystalline alumino-silicate frameworks comprising $[SiO_4]^{4-}$ and $[AlO_4]^{5-}$ tetrahedral units. T atoms (Si,Al) are joined by oxygen bridges. Introduction of an overall negative surface charge requires counterions (e.g., Na^+, K^+ and Ca^{2+}).

Synthetically they are made by a crystallization process with use of a template (e.g., tetrapropyl ammonium) that builds a self-assembling skeleton together with the silicon, aluminum, and oxygen atoms. Afterward the template is removed (e.g., by heating, the calcination process) leaving small nanosized holes and tubes along the crystal planes of the zeolite structure. The framework structure may contain linked cages, cavities, or channels, which are of the right size to allow small molecules to enter (i.e., the limiting pore sizes are roughly between 3 and 10 Å in diameter). In all, over 130 different framework structures are now known.

Zeolites have the ability to act as catalysts for chemical reactions which take place within the internal cavities. An important class of reactions is that the class is catalyzed by hydrogen-exchanged zeolites, whose framework-bound protons give rise to very high acidity. This is exploited in many organic reactions, including crude oil cracking, isomerization, and fuel synthesis. Zeolites can also serve as oxidation or reduction catalysts, often after metals have been introduced into the framework. Examples are the use of titanium ZSM-5 in the production of caprolactam, and copper zeolites in NO_x decomposition. See Figure 3.

Top layers of zeolites have also been deposited on many microporous support materials (e.g., alumina) to study their gas selective properties or to separate branched monomers (e.g., alkenes) from unbranched ones.

Table 1. Some membrane separation processes arranged according to the mechanism of separation.

Separation mechanism	Membrane separation process
Size exclusion (filtration)	nanofiltration, ultrafiltration, microfiltration
Solubility/diffusivity	reverse osmosis, gas separation, pervaporation, liquid membranes
Charge	electrodialysis

Table 2. Filtration processes with their properties and application.

Filtration process	Pore size	Separation capability	Pressure (bar)	Application
Molecular sieving	0.3–1 nm	branched molecules versus unbranched molecules	0–3	gas separation with zeolite molecular sieves
Nanofiltration	1–10 nm	permeation of low molecular (200–20,000 daltons) substances	5–25	purification of sugar from acids and salts weight from dyes, water treatment
Ultrafiltration	5–100 nm	retention of viruses, bacteria, dissolved substances with molecular weight between 10,000 and 500,000 daltons	0.5–5	dairy industry, beverage industry, pharmaceutical industry, separation of water from crude oil, separation of fruit and vegetable extracts, waste water treatment
Microfiltration	50 nm–5 μm	retention of bacteria, Colloids, protozoa (*Cryptosporidium Legnonella*)	0.5–3 0.5–3	prefiltration in water treatment, dye industry, beverage clarification, removal of bacteria, etc.

Nanofiltration Nanofiltration is a relatively young description for filtration processes using membranes with a pore size ranging from 1 to 10 nm. This term has been introduced to indicate a specific domain of membrane technology in between ultrafiltration and reverse osmosis. One of the first researchers using the term "nanofiltration" was Eriksson in 1988 [6]. But already some years earlier a company called FilmTec started to use this term for their NF50 membrane which was supposed to be a very loose reverse osmosis membrane or a very tight ultrafiltration membrane [7]. See Figure 4.

But the history of nanofiltration dates back further in time. In 1970 Cadotte et al. already showed that reverse osmosis membranes, a nanofiltration membrane with slightly smaller pore size, could be made out of polyethyleneimine and toluenediisocyanate, forming a polyurea [8]. This membrane, the NS l00, was different in several aspects from the then-existing reverse osmosis membranes. It was the first reverse osmosis membrane not made by a phase inversion process using cellulose acetate or polyamide plus it showed a high salt retention.

All ceramic and some polymeric nanofiltration membrane are considered to be nanoporous, and most of the polymeric nanofiltration membranes are regarded as "dense," which means that no fixed pores are present in the membrane, but they contain a network structure that can be charged or uncharged.

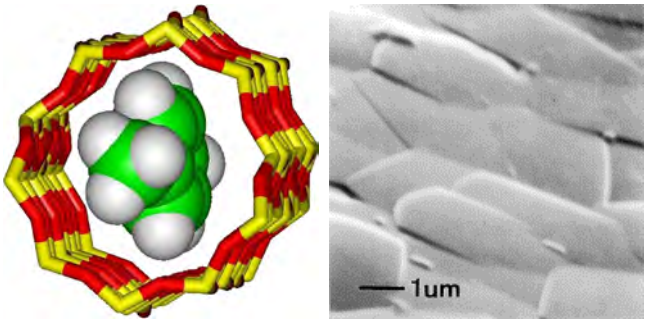


Figure 3. Left: Paraxylene molecule can diffuse freely in the open channels of silicalite. Right: top view of MFI-type ordered zeolite (molecular sieve) crystals having orthogonal nanopores of 0.55 nm on a silicon substrate. Courtesy of TOCK, TU Delft.

In several transport models for nanofiltration the pore size is used as a parameter to describe the morphology of the membrane [9, 10]. However, this concept of pores may be considered as hypothetical for several polymeric nanofiltration membranes. Bowen and Mukthar [11] mentioned that the determination of an effective pore size by transport models should not mean that those pores really exist in nanofiltration membranes. The hindrance to transport is the same for ions passing through the polymer network of a specific membrane as for ions passing through pores having these effective sizes.

In general nanofiltration membranes are used to separate relatively small organic compounds and (multivalent) ions from a solvent. However, in most of the applications with aqueous solutions, such as waste water, drinking water, and process water, the transport mechanisms are not yet fully understood. In the case of separation of organic solvents new developments of polymeric membranes are continuously showing up, where the solvent stability of these polymeric membranes is frequently studied and improved. These polymeric membranes may suffer from swelling when wrongly chosen.

Typical applications of these membranes are the separation of salts from dye solutions or the separation of acids from sugar solutions for the extraction of the purest products in highly concentrated form.

Ultrafiltration and Microfiltration Ultrafiltration membranes are porous membranes (in the range of 5 to 50 nm) that have a distinct, permanent porous network through

Figure 4. Schematic drawing of separation processes that occur in nanofiltration membranes.

which transport occurs [12, 13]. The separation of micro-organisms, bacteria, and solids from the liquid is mainly based on size and is pressure driven. The term ultrafiltration has been introduced to discriminate the process from microfiltration with a more narrow retention behavior. The membranes, in most cases polymeric, are either used in dead-end mode or in cross-flow mode. In dead-end mode all liquid is forced to pass the membrane, where in cross-flow mode, a tangential flow across the membrane is used to minimize a fouling layer, thereby keeping the filtration process ongoing [14]. See Figures 5 and 6.

For the retention of larger particles microporous membranes (with pore size from 50 nm to 5 μm) can be used. As an example, a feed containing macromolecules (e.g., proteins) is contacted with a membrane that contains small pores. The particles to be separated (the macromolecules) are withheld from flowing through the pores; the solution itself can flow through the pores. The reason that the particles cannot freely enter the pores is their geometric size, although interaction with the pore may also be of importance. In this case the properties of the membrane are also dependent on the structure of the membrane, together with the intrinsic properties of the membrane material. Separation of larger particles (e.g., micro-organisms) usually proceeds according to the same mechanism.

1.2.2. Separation by Solution–Diffusion Process

To achieve separation on a molecular scale a relatively dense membrane is required. In this membrane the separation of components takes place based on differences in diffusivity and differences in solubility. At the feed side, the components are being dissolved in the membrane and transported by diffusion through the membrane with a driving force acting inside the membrane. At the permeate side, the components leave the membrane. Separation results from differences in the solubility of the components into the membrane material and differences in velocities of diffusion through the membrane. The driving force is solely caused by properties of the membrane material, like affinity, and not of porosity (the geometry) of the membrane. Because the transport is a result of the solubility and the diffusivity, the separation is being called the "solution–diffusion" mechanism.

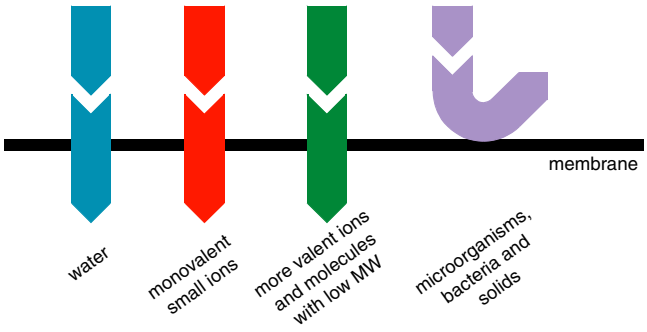


Figure 5. Schematic drawing of separation processes that occur in ultra- and microfiltration membranes.

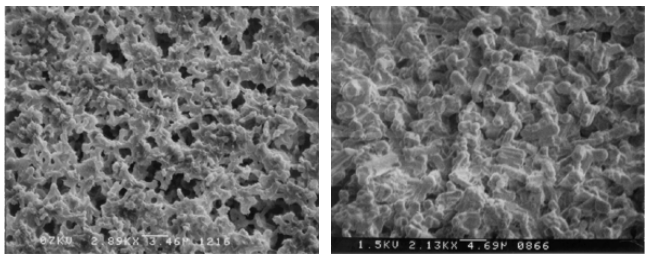


Figure 6. SEM micrographs of synthetic membrane filters. Left: Organic phase separated membrane filter. Right: Ceramic sintered filter. Courtesy of Aquamarijn Research.

Three main separation processes will be considered, reverse osmosis, gas separation, and pervaporation.

Reverse Osmosis In the reverse osmosis process, the processing liquid is being transported through the membrane under high pressure [15]. The membrane retains most of the ions in the liquid and larger solids and the process is often used to produce high quality water. Polymeric reverse osmosis membranes have no macroscopic pore structure but consist of a polymer network in which solutes can be dissolved. Like the membranes used for dialysis purposes, the membrane for reverse osmosis usually consists of a material that swells due to the processing liquid (e.g., water). This solvent-swollen network may be considered as a porous system comparable to dialysis membranes, although the structure of the latter is more loose than that of reverse osmosis membranes. It is difficult to distinguish single pores, as these are not fixed in place or in time due to the flexibility of the polymer chains

The distinction between dense and porous membranes is not very discrete. The membrane structure during the separation process then consists for a considerable part of volume that is filled with the feed. Membranes for dialysis purposes can easily consist for 50% of its volume of water. This volume may be called the pore volume, although the membrane matrix does not contain any distinct pores.

Gas Separation Gas separation (for example, separation of acid gases from natural gas or hydrogen in production processes) with polymeric membranes has become a widely known process for the last three decades. However, already over more than 150 years ago it was known that rubber balloons filled with gas lost their gas due to diffusion through the balloon wall [16, 17]. The real novel concept for gas separation using membranes was introduced by Monsanto [18]. Their membranes were able to recover hydrogen gas from the ammonia production.

For gas separation membranes it is difficult to achieve a good selectivity. The separation is being performed under harsh conditions and the membrane stability suffers under the presence of the process gas [19, 20]. For instance, CO_2 is a plasticizer (softener) for the polymer material, thereby weakening the membrane structure and thus decreasing the selectivity [21].

The preferred membrane for gas separation consists of a thin (<1 μm) top layer, supported by a rigid support with negligible transport resistance [22]. The thin top layer, the permselective layer, determines the overall gas flux. See Figure 7.

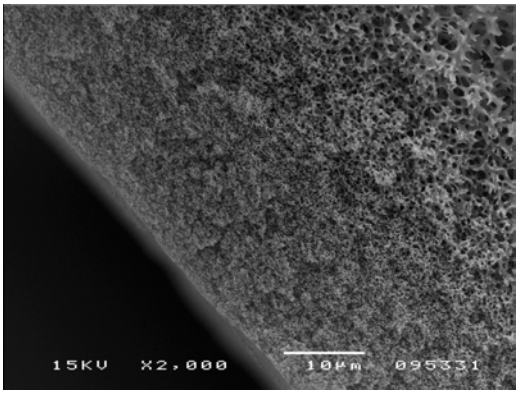


Figure 7. SEM cross-section of an asymmetric polymeric gas separation membrane with a gas selective dense skinlayer. Courtesy of Aquamarijn Research/Membrane Technology Group, University of Twente.

Membranes can be classified on the basis of their morphology or the separation process [18]. Current gas separation membranes are thin dense films, integrally skinned asymmetric membranes, or composites mainly prepared from glassy polymers [7]. Asymmetric membranes have a dense top layer and a porous substructure and are formed by a phase inversion process [19]. Composites have a dense top layer and a porous substructure. The top layer is created in a separate step for example by coating. In both cases, the permselective top layer should be as thin as possible to achieve a high flux. The substructure should have good mechanical strength with negligible gas transport resistance. Thin polymeric films by themselves are too weak to withstand the high differential gas pressures required in gas separation operations. Membranes with a support layer are therefore the most applied ones. The advantage of a composite membrane is that the top layer and the support can be optimized separately.

Pervaporation Pervaporation is a membrane separation process that can be used for the separation of liquid–liquid mixtures with an azeotropic composition or mixtures of components with a relatively small difference in volatility (dehydration of ethanol, acetic acid; removal of ethanol from a fermentation product). The transport occurs according to the solution–diffusion model through the nonporous membrane [23]. The diffusion through the membrane is in general enhanced by a lower partial pressure on the permeate side of the membrane (vacuum or a pure inert gas). Because the process uses low energy it is an advantageous alternative for conventional techniques like distillation.

The mixture flows through modules, where it passes along the membranes. The component to be removed is vaporized through the membranes and collected at very low pressure in a vacuum vessel. It is then condensed and purged out. A vacuum pump extracts the noncondensables and maintains the required vacuum. This vaporization cools down the processed mixture, which must be reheated to maintain the highest flux through membranes. Main applications are wine and beer dealcoholization, removal of organic solvents from aqueous streams, aroma recovery and concentration, and waste water purification.

1.2.3. Separation by Charge

In charge separation processes (membrane electrodialysis4), molecules in a liquid stream are being split in a positively charged and in a negatively charged ion. Using a combination of an anionic and a cationic membrane, water can be split in H^+ ions and OH^- ions. Combining this with liquids containing salts, direct acids and alkaline solutions can be made. In industrial applications this gives the opportunity for direct control of the pH of a process stream (juice industry and fermentation reactors) [24–26].

The anion exchange membranes carry fixed positive charges; the cation exchange membranes carry fixed negative charges. The presence of mobile counterions compensates these electrical charges: they carry a charge opposite that of the fixed charge, whereas the counterions are transported in the applied electric field.

A special type of ion exchange membrane is the bipolar membrane. It contains both the anionic membrane and the cationic membrane, but one of the membranes is permeable only for anions and the other only for cations. So there is no transport through the membrane possible. At the surface where the two types of membranes meet, water is split into OH^- and H^+ and these ions are transported out of the membrane. By this process no gases are being formed. Electrodialysis with bipolar membranes can be used to replace electrolysis of water, but with a wider variety of applications. Of course care should be taken that the two streams are separated from each other to prevent recombination into water [5, 27, 28]. See Figure 8.

The replacement of electrolysis by membrane electrodialysis seemed to be very promising two decades ago, but it was found that it is difficult to meet the specifications on a large scale. The available membranes did not have the ideal membrane requirements like a very high selectivity (the bipolar membrane layers are also permeable to salt co-ions), very high membrane permeability (yet the membrane layers add extra membrane resistance), and good long term stability

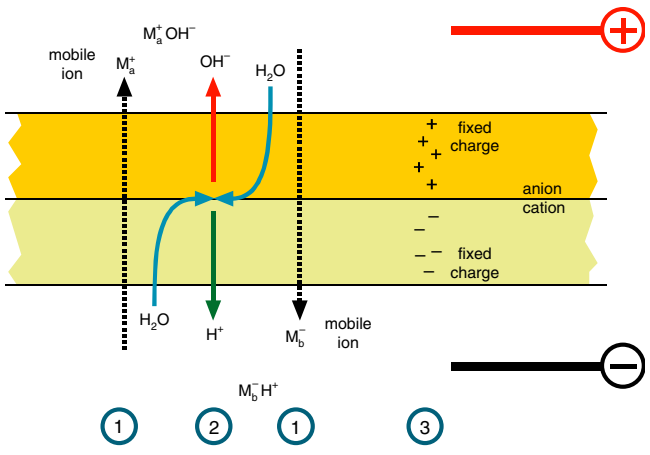


Figure 8. Processes that occur in and around an ion exchange membrane. In an externally applied electrical field (3) water is split in the membrane and the water ions are being transported (2) out of the membrane. If the membrane is permeable for ions (1), positively charged ions will move through the membrane toward the OH^- ions. Negatively charged ions will move toward the H^+ ions.

(the membranes are being attacked by a high concentration of alkaline, being produced in electrodialysis).

1.3. Membrane Structures

The structure of a membrane is vital for the performance of the membrane. The membranes are either symmetrical, where the properties of the membrane do not change through the cross-section of the membrane or they are asymmetrical (also known as composite). In the latter case the membrane is composed of a thin selective layer and a strong support layer. With some techniques it is possible to create both membrane and support layers in one single preparation step.

Asymmetric membranes are advantageous here compared to symmetric membranes because the flux determining top layer is thin, while the sublayer is responsible for the mechanical strength. A schematic representation of symmetric and asymmetric membranes is given in Figure 9.

Asymmetric membranes offer great possibilities in optimizing the membrane separation properties by varying the preparation parameters of especially the thin top layer. Also the ongoing development of polymers throughout the decades, like polysulfone (1965), polyether-ether-ketone (1980), and polyetherimide (1982), offers product engineers more and more possibilities for new membrane types.

1.3.1. Ceramic Membranes

Most ceramic membranes have an asymmetrical membrane structure with either a dense or a porous skin layer. The rough porous support is made of sintered ceramic particles [alumina (Al_2O_3), titania (TiO_2), and zirconia (ZrO_2)] that are gradually (in three to five steps) decreased in size before the top layer is applied. Typically the final layer of the support has a pore size between 1 and 5 μm. The preferred shape of ceramic membranes is a rod, because flat discs have shown to be too brittle.

Building the ceramic rod starts with making the support by injection molding the inner core with largest particle size. As long as the rod is not yet sintered (i.e., green phase), the paste of ceramic materials can be shaped freely. The rod is prebaked (fired) and subsequently a finer particle coating is applied by dip coating. This process of dip coating and firing is repeated several times until the final support coating is applied.

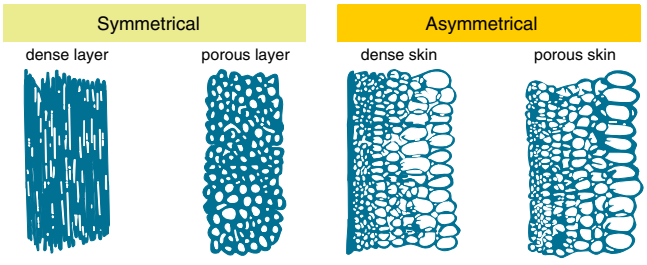


Figure 9. Schematic drawing of different membrane types. The membranes can be made as a fiber (circular) or as a flat sheet.

In the last step the skin layer is being applied, often again by several steps of dip coating, giving the membrane its separation properties. The whole stack is sintered for final fixation of all particles and ceramic substances. See Figure 10.

A typical α-alumina tube has a pore size of to 110–180 nm and is good for use as a microfiltration or an ultrafiltration membrane. For a finer membrane a γ-alumina layer can be applied on top of the α-alumina layer. By means of dip coating of a sol–gel, for example, from a Boehmite sol (γ-AlOOH) [29], a thin layer can be formed on top of the α-alumina. This layer is subsequently dried and heated up to 700 °C. By adding aluminum isopropoxide [Al(iso-$OHC_3H_7)_3$] in distilled water, a Boehmite sol can already be formed at 80 °C. Every dip coating and subsequent controlled drying step forms a layer with a thickness of around 0.5 μm [30]. During the drying the alumina solidifies, yielding a porous shell of alumina. The shell is then being hardened (calcined) by heating for several hours at high temperatures. The duration and temperature determine the final pore size, which for a γ-alumina layer is between 3 and 7 nm.

The grain size has also influence on the mechanical and physical properties of the ceramics: with decreased grain size, the membranes sized grains are stronger, are less brittle, and have higher temperature resistance [31, 32].

For certain applications an even finer pore size is wanted, for instance in a molecular sieve (separating gases, e.g., H_2/CH_4, H_2O/CH_4). In a chemical vapor deposition process [tetraethoxysilane and methyltriethoxysilane] an ultrathin silica can be applied that decreases the pore size even more. By carefully adjusting the gas flows through the rod and process temperatures, voids in the γ-alumina layer can be blocked, as the silica layer prefers to grow inside the voids. This is beneficial for the separation behavior of the silica layer, as the selectivity of the membrane increases [30]. The silica layer can also be applied in a multilayer sol–gel deposition [33, 34] instead of having an intermediate γ-alumina layer.

The process of sintering of particles can be applied to ceramics and polymers. Sintered polymeric membranes usually are macroporous membranes (pore size > 50 nm) but have after sintering the advantageous properties of ceramic membranes. See Figure 11.

Advantageous Properties of Ceramic Membranes

- Ceramics do not absorb water and do not swell. Swelling is a common problem with many materials, because when the membrane material absorbs water, the pore size increases, resulting in less retention.

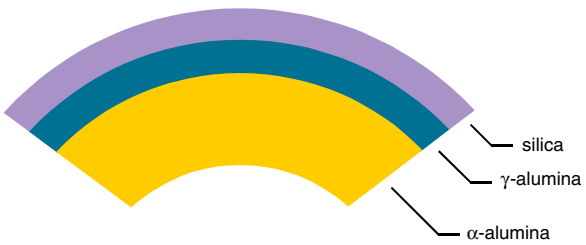


Figure 10. Schematic view of the different layers in a ceramic rod.

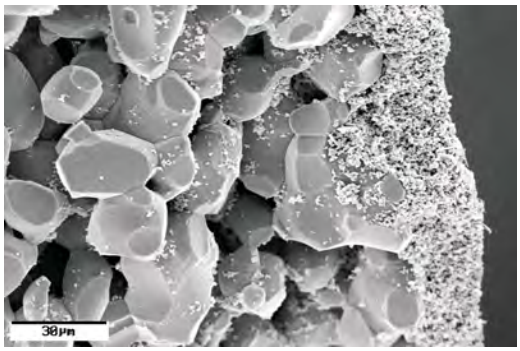


Figure 11. SEM photograph of a sintered ceramic membrane with a pore size of 800 nm. Courtesy of University of Karlsruhe.

- Ceramics are thermally stable. The membranes allow processes to be run at high temperatures. This makes the filtration for viscous fluids like oils easier, because their viscosity decreases.
- Ceramics are wear resistant (physically hard). This allows harsh cross-flow conditions beneficial for the removal of particles or cake layer, without damaging the membrane.
- Ceramics are chemically resistant. Most chemicals can be filtered or used during cleaning without attacking the membrane.

1.3.2. Polymeric Membranes

Due to the variety in properties of polymers and the processes to obtain these polymers, over the years many processes have been developed to obtain porous membranes.

Template Leaching Mixing two different components and molding this mixture into a certain heterogeneous structure, after which one of the components is removed by selective leaching, is a process used in the preparation of porous glass membranes.

Stretching of a Polymeric Film Dense polymeric films or foils can be stretched thereby generating voids in the film. By this method microfiltration membranes are formed. It is a technique that is commonly used for, for example, semipermeable membranes (for outdoor breathable clothing) with stretching a hydrophobic foil (for example PTFE). Best results are obtained with crystallizable polymers.

Phase Inversion A common technique to prepare the asymmetric membranes is phase inversion. While the process yields membranes in the whole spectrum from microfiltration to gas filtration, the membranes can be coated afterward to obtain other separation characteristics (by coating or interfacial polymerization).

Phase inversion (introduced by Cestings) is a very versatile technique. In one single process step various types of membranes can be made that have a porous support with or without a skin layer, where the skin layer can be dense or porous with a well defined pore size.

Phase separation is the process of changing a one-phase casting solution into at least two separate phases. The casting solution is uniform in chemical composition and in physical structure throughout the material. In all phase separation processes for membrane manufacturing, this casting solution is split into at least two phases: a solid material-rich phase that forms the porous/nonporous structure of the membrane and a material-poor phase that will be removed from the membrane. See Figure 12.

The phase separation is generally a thermodynamically driven process. After casting the solution onto a flat glass plate (or rotating cylinder in a continuous process) the formation of the pores starts by changing the composition of the solution, the temperature, or the pressure. The composition of the casting solution may be changed by bringing the casting solution in contact with a liquid not well miscible with the material. By changing the solution, the polymer solidifies and the pores are created.

Not only by immersion of the casting fluid in a nonsolvent can the solidification be started. It can also occur by evaporating a solvent from the casting solution or by reaction of components in the casting solution resulting in a nonsolvent agent.

Because for every polymeric material there is a combination of solvents/nonsolvents possible, the process can be used to produce a wide variety of membrane materials.

After solidification the polymers may be cross-linked or cured to improve functional properties (like strength and chemical resistance).

The types of phase separation processes are classified into six categories:

- Vapor-induced phase separation (VIPS), which is also known as “dry casting” or “air casting.” The evaporation of the solvent in the casting solution will result in a dense or porous product depending on the material and the used solvent/solute mixtures. The vapor may contain water or organic solvent molecules that may be absorbed by the casted film and will influence the porosity of the product.
- Liquid-induced phase separation, mostly known as “immersion casting” or “wet casting”: This process normally yields dense or porous products depending on the material and the solvent and nonsolvent mixtures used. This technique may be combined with VIPS for polymer solutions with, for example, two solvents with different boiling points [35, 36].
- Thermally-induced phase separation, frequently called “melt casting”: The material may solidify to a dense or porous product of a material-solvent-additive mixture

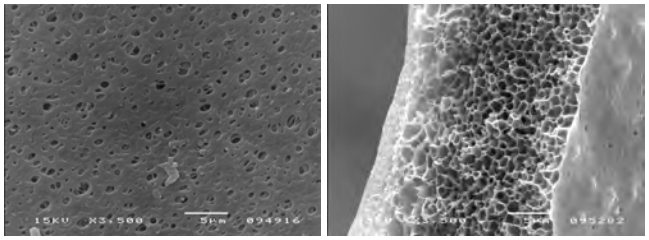


Figure 12. SEM photographs of liquid induced phase separated polymeric membranes. Left: top view skin layer. Right: cross-section showing a relatively dense skin layer on the retentate side and an open skin layer at the permeate side. Courtesy of Aquamarijn Research/Membrane Technology Group, University of Twente.

by varying the temperature of the material-solvent during the casting process.
- Colloidal-induced phase separation: A phase separated colloidal solution is used to perform structure arrestment on the microfabricated mold. For example, a polymerizable bicontinuous microemulsion can be applied onto the mold or a sol/gel mixture of a ceramic material. The pore morphology can be tuned by the addition of a co-solvent resulting generally into smaller pores.
- Pressure-induced phase separation, "pressure casting": The casting solution may contain for example a saturated dissolved gas. Reduction of the pressure (or increase of temperature) will induce growth of gas cells in the casting solution with a closed cell or open cell morphologies and a typical size of 0.01–1000 micrometers.
- Reaction-induced phase separation: A casting solution containing monomers starts to react and initiate phase separation due to for instance increase in molecular weight or production of a nonsolvent.

Coating of Membranes Existing (porous, polymeric) membranes are often coated to prepare composite membranes. On top of an asymmetric, porous membrane a thin layer of a polymeric solution is deposited resulting in a membrane with separation properties determined by the thin layer.

A second possibility is coating the complete membrane by a polymer that is not influencing the separation characteristics of the skin layer of the asymmetric membrane (a nonselective polymer). By this the voids in the asymmetric membrane can be plugged and the selectivity of the original membrane is increased.

Interfacial Polymerization A coating can also be applied by interfacial polymerization. In this technique the pores in the membrane are first filled with a liquid A. After immersing the membrane in a bath containing a reactant for liquid A, a polymer is being formed at the interface between liquid A and the bath. By this a new selective layer is being formed in the pores of the membrane.

1.3.3. Nanocomposite Membranes

There are a large number of nanofiltration and gas separation applications with membranes made from a polymeric or an inorganic material. The application of, for example, a polymeric material for a separation membrane depends of course upon both the throughput and the purity of the product transported through the membrane. This means that the permeability coefficient and the selectivity should be as large as possible. However, it has been found that simple structural modifications, which lead to an increase in product flux, usually cause a loss in permselectivity and vice versa [37–41]. This so-called "trade-off" relationship is well described in the literature [42, 43]. In recent years, the efforts and successes in synthesizing a variety of nanostructured hybrid materials have provided a new degree of freedom for the development of advanced materials with enhanced separation properties [44]. See Figure 13.

Nanocomposites represent a current trend in developing novel nanostructured materials. They can be defined as a combination of two or more phases containing different compositions or structures, where at least one of the phases is in the nanoscale regime. Kusakabe and his co-workers [45] already reported in 1996 that the permeability of CO_2 in a polyimide/SiO_2 hybrid nanocomposite membrane is 10 times larger than in the corresponding polyimide.

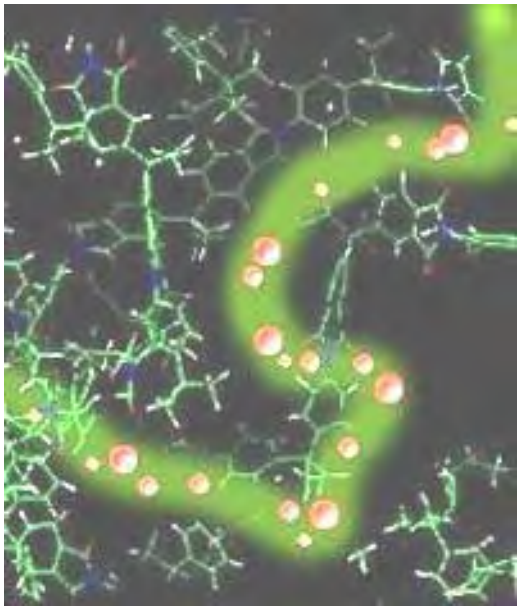


Figure 13. Gas molecules (red) move through the spaces between the polymer chains (in green). The larger the spaces, the faster and more selective the movement. The finding, which appeared in *Science*, describes a new type of nanoparticle-enhanced filter for separating compounds at the molecular level. Reprinted with permission from [46], A. Hill et al., *Science* 296, 519 (2002). © 2002, American Association for the Advancement of Science.

Another application is the use of zeolite nanocrystals (10–100 nm) to combine the advantages of polymers and zeolites while overcoming the shortcomings of both. Polymer–zeolite nanocomposite membranes can be developed for air separation as a promising alternative to the conventional energy-intensive cryogenic distillation. Utilization of polymer–zeolite nanocomposite membranes shows a route to achieve a modified polymer matrix offering high molecular sieving selectivities ($O_2/N_2 > 20$) while maintaining polymer processing conditions.

For any molecule to move across a membrane, it must go through a solution–diffusion process. The molecule first has to get into the membrane—the solution part of the process—then diffuse through it. Recently novel gas separation nanomembranes have been obtained with inorganic silica nanoparticles, with contradictory, but very useful, properties. It was shown that larger molecules dissolved much faster into the membrane, and once in, moved right through to the other side before the smaller molecules had completed the first step.

Silica nanoparticles embedded in a carbon base may help in the future to produce gases free of impurities. Because of the nanocomposite's ability to trap molecular-sized impurities it could be further used in processes such as biomolecule

purification, environmental remediation, sea water desalination, and petroleum chemicals and fuel production [46].

1.3.4. Other Membrane Structures

Track Etched One of the most successful attempts to manufacture the best membrane (a sieve) is the track-etched membraney. For the production of track-etched membranes, dense polymeric films (polycarbonate or polyester) are randomly exposed to a high-energy ion bombardment or particle radiation. This bombardment damages the polymeric chain in the dense film, leaving small "tracks." Subsequently the polymer can be etched specifically (in an acid or alkaline solution) at the ends of the damaged track. The pores that are being formed are cylindrical channels and very uniform in size. However, the membrane has a low porosity, due to the fact that the chance for an overlap between two pores increases with the porosity (the density of the radiation). The membranes are often used in laboratories for analysis. See Figure 14.

Anodized Alumina Membranes The formation of the pores of an anodized alumina membrane is a self-assembling process. By anodization of aluminum in an acidic solution, a highly ordered structure of pores in the Al_2O_3 matrix can be obtained. Due to lattice expansion by the oxidation of the aluminum, an anisotropic potential distribution, and heat development during anodization, the self-organization process forms the pores. The membranes are relatively thick, resulting in long pores with a pore size ranging from 20 to 200 nm. The pore size is very uniform (see also Fig. 15). However, the membranes are unsupported and need, depending on the application, a second support.

Microsieve Most of the previously described porous membranes have a spongelike structure, which makes it difficult to clean the inside of the membranes. Also the retention of particles, proteins, and other coagulates is not only determined by the porous top layer, but also by the sponge support structure. Moreover, in certain applications, like clarification of beverages, the cleaning agents must be removed completely after cleaning to prevent contamination of the beverage. However, it is nearly impossible to remove cleaning agents that are left behind in dead-end pores. Until recently the best sieve structures were the track-etched and

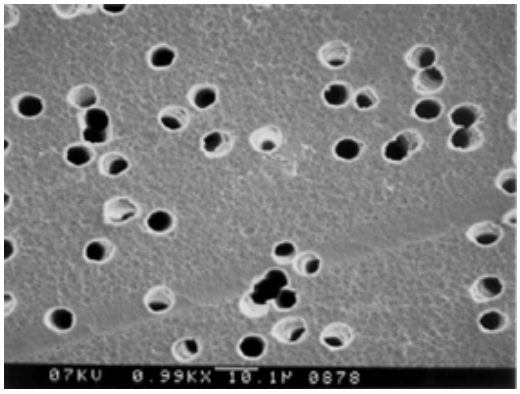


Figure 14. SEM micrograph of polymeric track etched membrane filter. Courtesy of Aquamarijn Research.

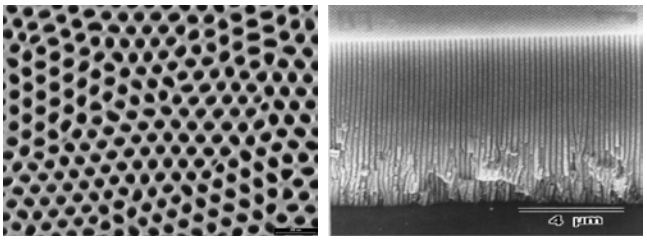


Figure 15. SEM photograph of an anopore membrane with a pore size of 50 nm. Left: top view. Right: side view. Note the vertical pore/channel structure. Reprinted with permission from [201], H. Asoh et al., *J. Electrochem. Soc.* 148, 152 (2001). © 2001, The Electrochemical Society.

the anopore membrane. However, both membrane types lack the possibility of adjusting the pore size freely in conjunction with the highest achievable porosity.

Using techniques used in semiconductor technology it is possible to fabricate thin membranes with perfectly uniform pores [3]. The membrane, the microsieve, was made using flat substrates, thin-layer deposition techniques, photolithography, and high-resolution etching methods.

2. NANO- AND MICROENGINEERED MEMBRANE TECHNOLOGY

Micro- and nanoengineering refers to the technology and practice of making three-dimensional structures and devices with dimensions on the order of micrometers, respectively nanometers. When looking at such small devices, a number of physical effects have different significance on the micrometer scale compared to macroscopic scales. This includes such topics as micromechanics, which deals with the moving parts of micro/nanoengineered devices, and micro/nanofluidics, which deals with the fluidic properties of such small devices.

The two constructional technologies of microengineering are microelectronics and micromachining. Microelectronics, producing electronic circuitry on silicon chips, is a very well developed technology. Micromachining is the name for the techniques used to shape the structures and moving parts of microengineered devices. In recent years micromachining and microengineering have become synonyms, also because of the advent of nanoengineering techniques (i.e., techniques capable of shaping structures below the 100 nm scale). See Figure 16.

The formation of micro- and nanopores in a thin but strong membrane structure and its potential applications is the key element in engineered membrane technology. Up to now relatively large pores ($\gg$1 micrometer) were made in thin foils or membranes with conventional microperforation methods (e.g., laser drilling, electroforming). With microengineering techniques originating from the semiconductor industry, it is relatively easy to downscale and form submicrometer pores (down to 500 nm) using conventional photolithographic methods, with, for example, contact masks and wafer steppers.

Many applications, especially filtration applications, prefer pores in a membrane with a low flow resistance (i.e., the length of the pore should be as small as possible). Therefore

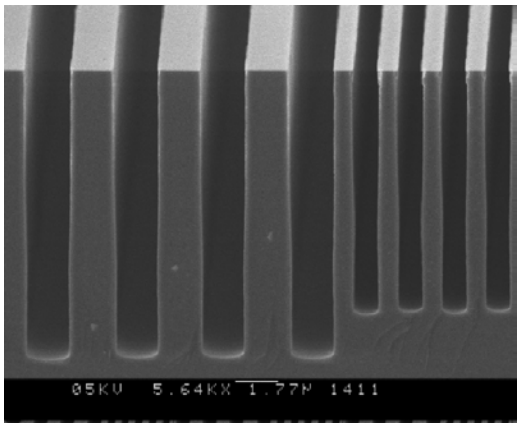


Figure 16. Micromachining: Deep trench etching in silicon. Reprinted with permission from [47], H. V. Jansen, Plasma Etching in Microtechnology, Thesis, 1996.

van Rijn and Elwenspoek [48] introduced the microsieve, a very thin membrane with a specially designed macroperforated support structure to strengthen the thin membrane. See Figures 17 and 18.

Using supported microsieve structures and submicrometer lithography techniques from the semiconductor industry it is relatively straightforward to produce microsieves with submicrometer features. The microengineered pores, which are well defined by photolithographic methods and anisotropic etching, allow, for example, accurate separation of particles by size. The membrane thickness is usually smaller than the pore size in order to keep the flow resistance small (one to three orders of magnitude smaller than other types of filtration membranes). See Figure 19.

Various lithographic process flows will now be discussed for economical production of microsieves with a high effective use of the membrane area. Also new nanopatterning techniques like laser interference lithography, shadow mask techniques (nanostencilling), and micro/nanocontact printing will be presented.

2.1. Micro- and Nanolithographic Patterning

2.1.1. Contact Mask Lithography

The contact mask technique makes use of a mask that is in direct contact with the photolacquer layer during illumination with a suitable UV light source. The sizes of the patterns in the mask are copied 1 to 1 to the size of the obtained patterns in the photolacquer layer. Practically contact masks can be used if the smallest feature sizes are larger than 0.5 micrometer, dependent on the wavelength of the used UV source.

Figure 20 shows how the contact mask technique is used to produce a microsieve with a specially designed support structure in a silicon ⟨100⟩ wafer. During the KOH etching process silicon is etched from the back side as well as from the front side through the pores of the silicon nitride membrane. See Figure 21.

A big disadvantage of using direct contact between the mask and the photolacquer layer is that the presence of small particles or contaminants (>1 micrometer!) between

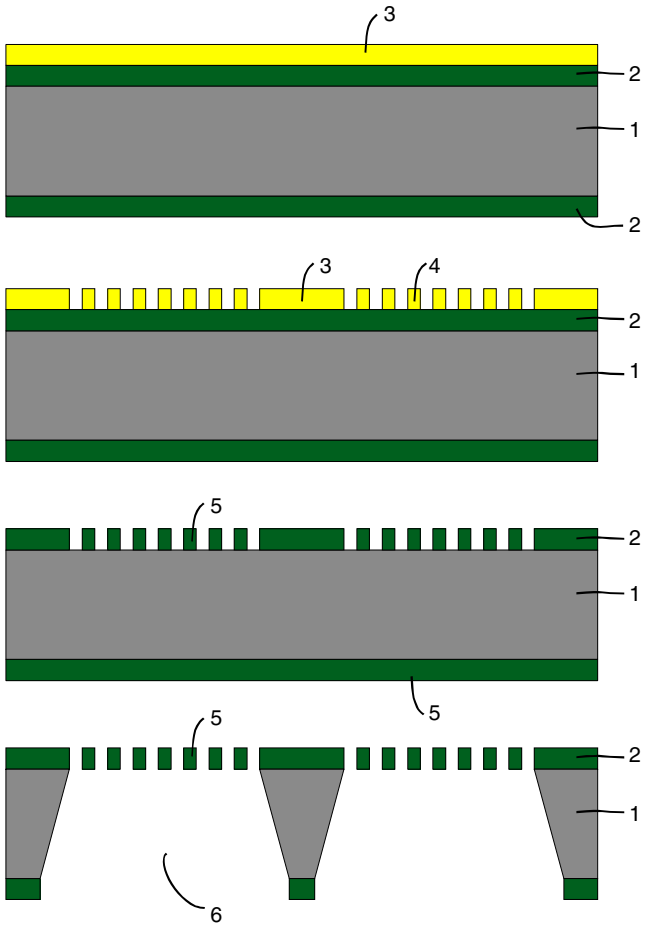


Figure 17. Basic process steps to produce a microsieve. Layer 2 is, for example, a low stress silicon nitride that is deposited on a 380 μm thick polished silicon wafer 1 by means of LPCVD (low-pressure chemical vapor deposition). On this 1 μm thick silicon nitride layer 2 a photosensitive lacquer layer 3 is formed by spin-coating. This layer is patterned with small pores 4 by exposing it to ultraviolet (UV) light through a photomask and subsequently developing it. The pattern in the photosensitive layers 3, 4 is transferred into the silicon nitride membrane 5 by means of RIE (reactive ion etching) with a CHF_3/O_2 plasma.

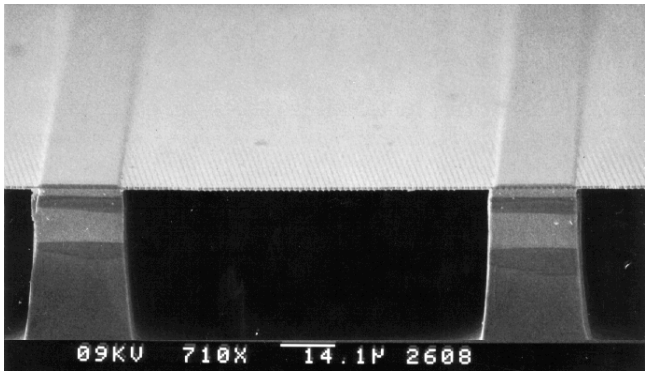


Figure 18. SEM micrograph of a microsieve with a membrane and a support structure. Courtesy of Aquamarijn Research.

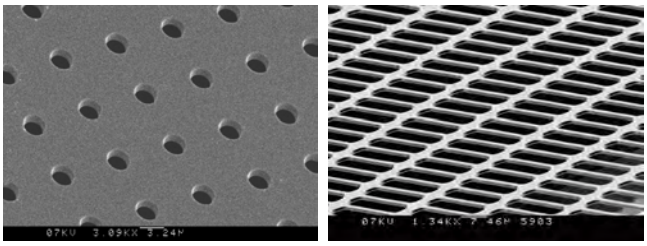


Figure 19. SEM micrograph of a microsieve with circular and slit-shaped pores. Courtesy of Aquamarijn Research.

the mask and the photolacquer layer may blur the pattern transfer. The amount of blurring is of course directly related to the ratio of the wavelength of the used light and the smallest feature size. Defect-free processing of wafer sized microsieves with small perforations is therefore virtually impossible even in high class clean room environments. A particle with a size of 10–20 micrometer between the

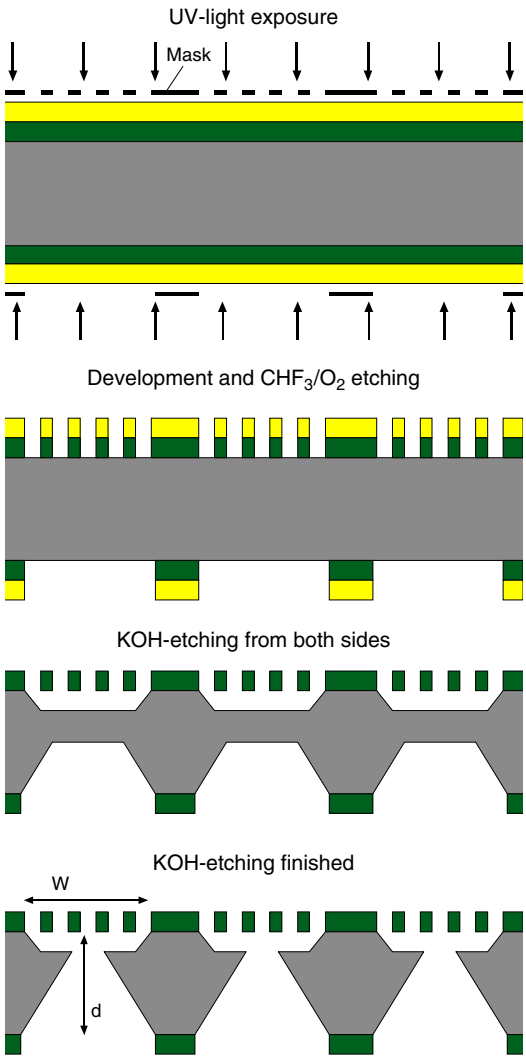


Figure 20. Fabrication process of a microsieve with contact mask lithography.

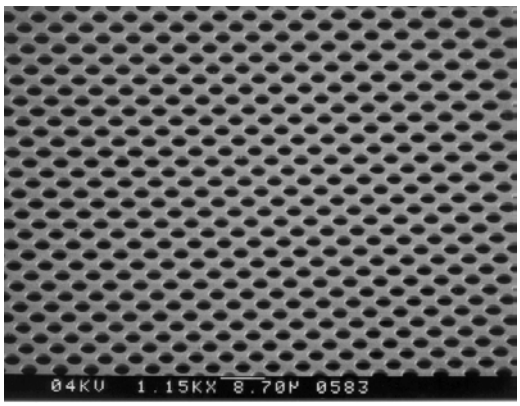


Figure 21. SEM micrograph of the surface of a microsieve where the silicon is etched and removed via the perforations in the membrane. Courtesy of Aquamarijn Research.

contact mask (e.g., patterned chromium layer on a glass plate with thickness 2 mm) and the lacquer layer is able to blur or destroy the pattern transfer over an wafer area larger than 1–100 mm^2. For wafer sized microsieves or for microsieves with a pore size smaller than 1 micrometer one often makes use of other more modern lithographic techniques, such as wafer steppers, to produce microsieves.

2.1.2. Wafer Stepper Lithography

In light projection lithography [49–51] an optical lens system is used to project the pattern of the mask onto the photolacquer layer (contact-free technique). The blurring of the pattern is now only restricted to the size of the particles present on the photolacquer layer. The projection system often reduces the pattern of the mask between 1 and 10 times with respect to the projected pattern in the lacquer layer. The typical lacquer area that can be projected in one step is about 1–2 cm^2. More steps (of the wafer stepper) are therefore needed to project the full wafer (3–12 inch diameter). A typical single step time is 1–2 seconds. This time cannot easily be reduced because of the mechanical repetition step time and the minimum exposure time per light shot. Although wafer steppers are very expensive (new generation >5 M euro) they are very economical in mass production applications. Also it is relatively more easy to get smaller pattern dimension (e.g., 0.12 micrometer linewidths) than contact mask applications due to the use of high resolution lenses with a large numerical aperture. See Figure 22.

Additional advantages are that the mask pattern is relatively large and can be made with relatively cheap laser writers.

Lithographic reduction has historically been accomplished by optimizing the parameters in the Rayleigh model for image resolution: In this model [53, 54], image resolution = k_1/NA, and depth of focus = $k_2\lambda/\mathrm{NA}^2$, where λ = exposure wavelength and NA = numerical aperture (k_1, k_2 are constants for a specific lithographic process). To pattern devices with decreasing feature sizes, photolacquer exposure wavelengths were reduced and numerical apertures were increased [55, 56]. See Table 3.

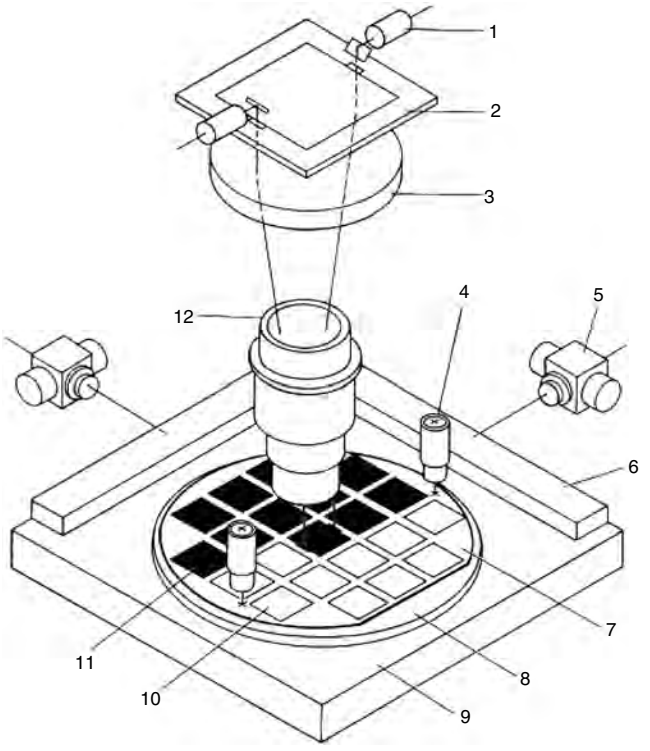


Figure 22. Wafer stepper. (1) Optical adjustment system, (2) reticule, (3) telecentric lens, (4) alignment optic, (5) laser interference meter, (6) mirror, (7) wafer, (8) wafer chuck, x–y table, (9, 10) projected and unprojected pattern fields, (12) objective. Reprinted with permission from [52], D. Widmann et al., "Technologie hochintegrierter Schaltungen (2. Auflage)." Springer-Verlag, Berlin/Heidelberg/New York, 1996. © 1996, Springer-Verlag.

2.1.3. E-beam and Future Semiconductor Lithography

Electron beam lithography (EBL) is a specialized technique [57, 58] for creating the extremely fine patterns required by the modern electronics industry for integrated circuits. Derived from the early scanning electron microscopes, the technique in brief consists of scanning a beam of electrons across a surface covered with a resist film sensitive to those electrons, thus depositing energy in the desired pattern in the resist film. The process of forming the beam of electrons and scanning it across a surface is very similar to what happens inside the everyday television or cathode ray tube display, but EBL typically has three orders of magnitude better resolution (typical spot size for EBL is 18 nm). The main attributes of the technology are (1) it is capable of very high resolution, almost reaching the atomic level; (2) it is a flexible technique that can work with a variety of materials and an almost infinite number of patterns; (3) it is slow, being one or more orders of magnitude slower than optical lithography; and (4) it is expensive and complicated—electron beam lithography tools can cost many millions of dollars and require frequent service to stay properly maintained.

The first electron beam lithography machines, based on the SEM, were developed in the late 1960s. Shortly thereafter came the discovery [59] that the common polymer PMMA (polymethyl methacrylate) made an excellent electron beam resist. It is remarkable that even today, despite sweeping technological advances, extensive development of commercial EBL, and a myriad of positive and negative tone resists, much work continues to be done with PMMA resist on converted SEMs.

Currently, electron beam lithography is used principally in support of the integrated circuit industry, where it has two niche markets. The first is in mask making, typically the chrome-on-glass masks used by optical lithography tools. It is the preferred technique for masks because of its flexibility in providing rapid turnaround of a finished part described only by a computer-aided design file. The ability to meet stringent linewidth control and pattern placement specifications, on the order of 20 nm each, is a remarkable achievement. See Table 4.

Second, EBL is used for research into the scaling limits of integrated circuits and studies of quantum effects and other novel physics phenomena at very small dimensions. Here the resolution of EBL makes it the tool of choice. A typical application is the study of the Aharanov–Bohm effect [60, 61], where electrons travelling along two different paths about a micrometer in length can interfere constructively or destructively, depending on the strength of an applied magnetic field. Other applications include devices to study ballistic electron effects, quantization of electron energy levels in very small structures, and single electron transistors [62, 63].

To see these effects typically requires minimum feature sizes of 100 nm or less as well as operation at cryogenic temperatures.

Alternative Techniques It is prudent to consider possible alternatives before committing to EBL technology. For chrome-on-glass optical mask fabrication, there are optical mask writers available that are based either on optical reduction of rectangular shapes formed by framing blades

Table 3. Deep UV and extreme UV monochromatic laser light sources.

Wavelength (nm)	Laser type
248	KrF
193	ArF
157	F_2

Table 4. New predictions (year node potential solutions) for the ever decreasing linewidths from the 1999 ITRS Lithography Roadmap.

Year	Linewidth	Lithographic tool
1999	180 nm	KrF
2002	130 nm	KrF + RET, ArF
2005	100 nm	ArF + RET, F2, EPL, PXL, IPL
2008	70 nm	F2 + RET, EPL, EUV, IPL, CBDW
2011	50 nm	EUV, EPL, IPL, EBDW
2014	35 nm	EUV, IPL, EPL, EBDW
>2014		innovative technology

Note: KrF, ArF, and F2 are laser types. RET is resolution enhancement technology, including phase-shift-mask technology for enhanced resolution. EPL is electron projection lithography. EUV is extreme ultraviolet. IPL is ion projection lithography. PXL is proximity X-ray lithography. EBDW stands for electron-beam direct-write techniques.

or by multiple individually controlled round laser beams. Although at present EBL is technologically ahead of optical mask writers, this may not continue in the future. However, EBL will continue to provide a resolution advantage over the optical mask writers which may be important for advanced masks using phase shift or optical proximity correction. For 1:1 mask fabrication (i.e., X-ray), EBL will continue to be the most attractive option.

Optical lithography using lenses that reduce a mask image onto a target (much like an enlarger in photography) is the technique used almost exclusively for all semiconductor integrated circuit manufacturing. Currently, the minimum feature sizes that are printed in production are a few tenths of a micrometer. For volume production, optical lithography is much cheaper than EBL, primarily because of the high throughput of the optical tools. However, if just a few samples are being made, the mask cost (a few thousand dollars) becomes excessive, and the use of EBL is justified. Today optical tools can print 0.12 μm features in development laboratories, and 0.08 μm should be possible within a few years.

By using tricks, optical lithography can be extended to 0.1 μm or even smaller. Some possible tricks include overexposing/overdeveloping, phase shift and phase edge masks, and edge shadowing [64]. The problem with these tricks is that they may not be capable of exposing arbitrary patterns, although they may be useful for making isolated transistor gates or other simple sparse patterns. Another specialized optical technique can be used to fabricate gratings with periods as small as 0.2 μm by interfering two laser beams at the surface of the sample. Again, the pattern choice is very restricted, although imaginative use of blockout and trim masks may allow for the fabrication of simple devices [65].

X-ray proximity printing may be a useful lithographic technique for sub-0.25 μm features [66]. Again, it requires a mask made by EBL, and since the mask is 1:1 this can be a formidable challenge. Especially if the throughput required exceeds the limited capabilities of EBL, this may be an attractive option. The disadvantage is that X-ray lithography is currently an extremely expensive proposition and the availability of good masks is limited. It also requires either a custom built X-ray source and stepper or access to a synchrotron storage ring to do the exposures. With care, X-ray lithography can also be extended to the sub-0.1 μm regime [67].

Another technique to be discussed is ion beam lithography. The resolution, throughput, cost, and complexity of ion beam systems is on par with EBL. There are a couple of disadvantages, namely, limits on the thickness of resist that can be exposed and possible damage to the sample from ion bombardment. One advantage of ion beam lithography is the lack of a proximity effect, which causes problems with linewidth control in EBL. Another advantage is the possibility of *in-situ* doping if the proper ion species are available and *in-situ* material removal by ion beam assisted etching. The main reason that ion beam lithography is not currently widely practiced is simply that the tools have not reached the same advanced stage of development as those of EBL.

2.1.4. Laser Interference Lithography

In 1995 van Rijn [3] proposed for the first time the use of laser interference lithography for the production of micro- and nanoengineered membranes (micro- and nanosieves). Because of the relatively simple periodic structures (orifices and slits) that are needed for the production of nanosieves and the fact that laser interference is potentially a very nonexpensive patterning technique [68–72] that is also applicable on nonplanar surfaces (very large focus depth) this method will be discussed more in detail, also because there is little published material on this subject so far [73]. In 1996 in a collaboration of van Rijn [74] with G. J. Veldhuis [127], A. Driessen, P. Lambeck, and Professor C. Lodder of the University of Twente an existing laser apparatus was modified to obtain photolithographic patterns for the development of nanosieves and isolated magnetic domains [76].

Double Laser Interference Exposure Technique When two planar waves of coherent light interfere, a pattern of parallel fringes will appear. These fringes can be used for the exposure of a photosensitive layer. See Figure 23.

The depth of focus of this method is dependent on the coherence length of the light and can be on the order of meters or more, compared to (sub)micrometers for conventional optical lithography systems. As a result the demands on substrate flatness and wafer positioning are less critical.

After the first exposure the substrate is rotated over 90° and exposed to laser interference lines again. Now the gratings cross each other and after development a square array of lacquer pores (Fig. 24, left) remains. The exposure time of the photolacquer layer is a critical factor. In case the exposure time is chosen larger the middle pattern of Figure 24 will be obtained. Upon further increasing this time isolated photolacquer dots (right) remain.

Device Fabrication with Relatively Short Exposure Time Part of an incoming plane wave is reflected by the mirror and interferes with the undisturbed part of the wave to form an interference pattern (grating) on the substrate surface. To produce the plane wave, TE polarized light of an argon laser with a wavelength $\lambda_{uv} = 351.1$ nm is spatially filtered and expanded by focusing it onto a pinhole. See Figure 25.

If the light intensity of each beam is I_0, the radiance on the surface is given by

$$I = 4I_0 \sin^2\left(\frac{\pi x}{\Lambda_x}\right)$$

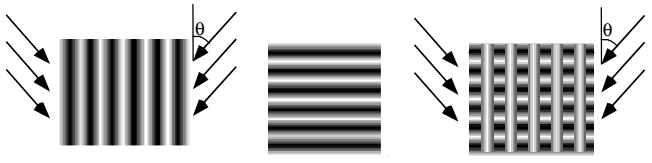


Figure 23. Left: first exposure of the substrate with the photolacquer layer. Middle: rotation of the substrate 90°. Right: second exposure of the substrate.

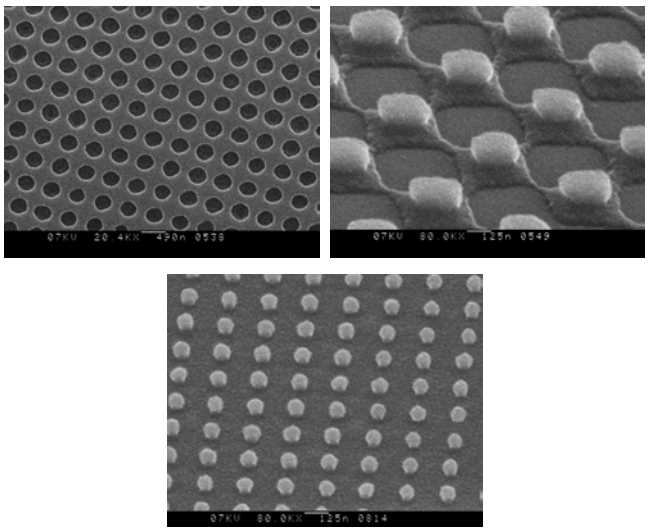


Figure 24. SEM micrograph of photolacquer layer that remains after a double exposure in the laser interference setup. Left: short exposure time, middle: intermediate exposure time. Left and middle: Courtesy of J. Micromech. Microeng. Right: long exposure time. Courtesy of Aquamarijn Research.

with Λ_x the fringe period in the x-direction planar to the photolacquer layer:

$$\Lambda_x = \frac{\lambda_{\mathrm{UV}}}{2 \sin \theta}$$

Here λ_{uv} is the wavelength of the laser light in the medium that surrounds the substrate (usually air) and θ is the half-angle between the two beams. The smallest period that can theoretically be obtained occurs for $\theta = 90°$ and is equal to $\lambda_{\mathrm{uv}}/2$.

The smallest period that can theoretically be obtained with our configuration is $\Lambda = \lambda_{\mathrm{uv}}/2 = 175$ nm. The corresponding minimum pore size (with a porosity >30%) will be approximately $175/2 = 88$ nm. At low porosity of course smaller pore sizesa can be made.

Since the beam is only split for a short path length near the substrate, this setup is very insensitive to mechanical instabilities and no feedback loop [77] is required to stabilize the interference pattern.

The thickness of the photosensitive layer needs to be chosen with care to avoid problems with the periodic pattern normal to the substrate surface due to interference between the incoming beam and the one reflected on the substrate surface. Its period is given by $\Lambda_\perp = \lambda_{v\varpi}/2n \cos\theta_n$ where n is the refractive index of the photoresist and $2\theta_n$ is the angle between the beams inside the resist. We used $\theta = 20°$ and with $n = 1.7$ at $\lambda_{v\varpi} = 351.1$ nm we find $\Lambda = 510$ nm and $\Lambda_\perp = 105$ nm. Therefore the thickness of the photoresist layer is chosen smaller than 105 nm.

The area that can be patterned using our mirror of 2.5×2.5 cm^2 equals approximately 9×9 mm^2 for $\Lambda = 510$ nm.

In Figure 26, the backside of a support (1), a single crystalline 3″ ⟨100⟩-silicon wafer with a thickness of 380 μm, is pre-etched to a thickness of 15 μm using optical lithography and conventional KOH etching (25%, 70 °C). On the front side of the pre-etched support (1) a layer (2) of amorphous low stress [78] silicon rich silicon nitride with thickness 0.1 μm is deposited by means of LPCVD by reaction of dichloresilane (SiH_2Cl_2) and ammonia (NH_3) at a temperature of 850 °C. Except at the area where the microsieve pattern will be formed an etch mask layer (3) of sputtered chromium with a thickness of 30 nm is deposited.

On top of this chromium layer (3) a layer (4) of positive resist was spun and patterned using interference lithography. A 100 nm thick layer (4) of positive photoresist (Shipley S1800 series) was spun, followed by a 5 min prebake at 90 °C to evaporate the solvent. The resist was exposed to the

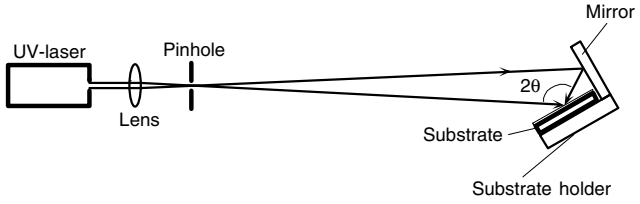


Figure 25. Setup for laser interference lithography. Courtesy of J. Micromech. Microeng.

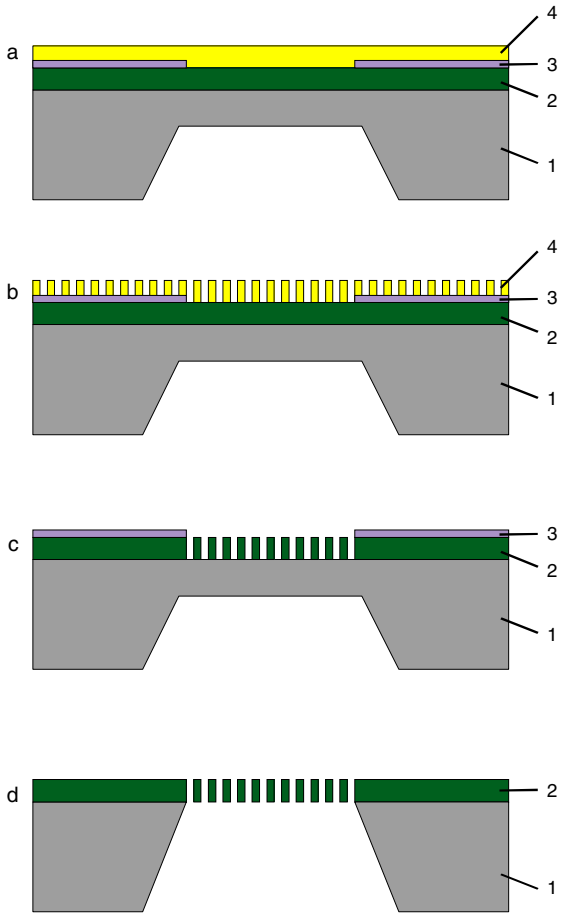


Figure 26. Schematic representation of the fabrication process of a microsieve. The numbers indicate the silicon support (1), the silicon nitride membrane (2), the chromium etch mask (3), and the photoresist layer (4).

interference line pattern for 45 s. The intensity of the incoming light in the exposed area was measured to be 2 mW/cm^2 for normal incidence ($\theta = 0°$). After rotating the substrate over 90° the exposure was repeated. The resist was developed for 15 s in a 1:7 mixture of Shipley–Microposit 351 developer and deionized water and dried by spinning. The exposure time was chosen such that only at the crossings of the grid lines (after first and second exposure) does the resist receive enough energy to be removed completely after development. Therefore a two-dimensional pattern of pores is created in the resist. A SEM picture of the exposed (2×45 s) and developed resist is given at the left in Figure 26. The diameter of the pores in the photoresist depends on the duration of exposure, herewith giving a possible tool to vary the pore size at a constant pore to pore distance. However, when the exposure time is chosen too long (2×75 s) the pores in the resist pattern may become too large and will overlap (Fig. 26, middle).

Next the interference pattern is transferred into the silicon nitride membrane layer (2) by means of CHF_3/O_2 reactive ion etching at 10 mTorr and 75 W for 2 minutes forming the required perforations. Subsequently the silicon underneath the membrane layer (2) is anisotropically etched with an SF_6O_2 plasma at 100 mTorr and 100 W for 10 minutes with an etch rate of 2 μm/min in order to form the macroscopic openings in the support (1). Figure 27 shows a SEM photograph of the resulting perforated membrane layer (2) showing a very regular pore pattern, the pore size being 260 nm with a pore to pore spacing of 510 nm. The pore size was very uniform over the whole 9×9 mm^2 area.

Microsieves with pore sizesa down to 65 nm have been fabricated using double-exposure laser interference lithography. The pores are obtained with an inverse process, as the direct process of pore formation in photolacquer has a narrow process latitude. An array of posts is transferred into an array of pores by evaporating chromium onto the posts, followed by a lift-off in acetone. The resulting patterned chromium layer is used as an etch mask for plasma etching of the silicon nitride membrane. See Figures 28 and 29.

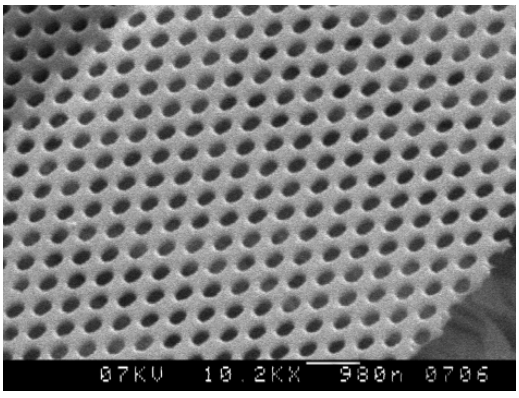


Figure 27. SEM micrograph of the microsieve membrane showing pores with a diameter of 260 nm in a 100 nm thick silicon nitride layer. Courtesy of Nanotechnology.

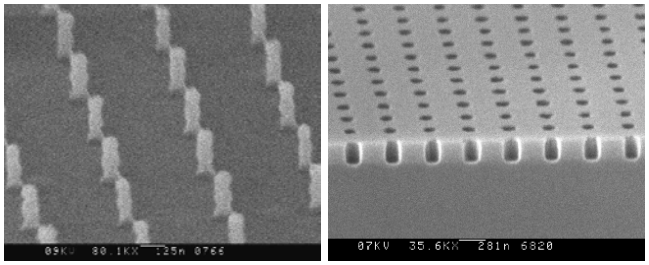


Figure 28. Left: SEM micrograph of 80 nm wide posts with rippled sidewalls caused by the vertical interference pattern. Right: SEM micrograph of a membrane after plasma etching through the pores in a chromium layer. Courtesy of Aquamarijn Research.

2.1.5. Nanocontact Printing and Etching with Phase Separated Membranes

Aloys Senefelder used in 1796 a porous stone (in Greek, *lithos*) as a tool for printing by patterning the stone with ink attracting (hydrophobic) and ink repelling (hydrophilic) regions. Lithography for semiconductor mass fabrication and other microsystem and nanotechnology applications has nowadays regained interest in inexpensive microprinting methods as an alternative or complement on current high tech optical wafer stepper technology. A need exists therefore in the art for a convenient, inexpensive, and reproducible method of plating or etching a surface according to a predetermined pattern. The method would ideally find use on planar or nonplanar surfaces and would result in patterns having features in the micrometer and submicrometer domain. Additionally, the method would ideally provide for convenient reproduction of existing patterns. Additionally, the need exists for the fabrication of surfaces that can pattern portions (e.g., SAMs) amenable to attachment of biological species, such as antibodies, antigens, proteins, cells, etc., on the (sub)micrometer scale.

The study of self-assembled monolayers (SAMs) is an area of significant scientific research. Such monolayers are typically formed of molecules each having a functional group that selectively attaches to a particular surface, the remainder of each molecule interacting with neighboring molecules in the monolayer to form a relatively ordered array. Such SAMs have been formed on a variety of products including metals, silicon dioxide, and gallium arsenide using relief printing with a molded stamp made from polydimethylsiloxane (PDMS) [79]. The upper relief part of the stamp provided with a suitable SAM coating is then being contacted

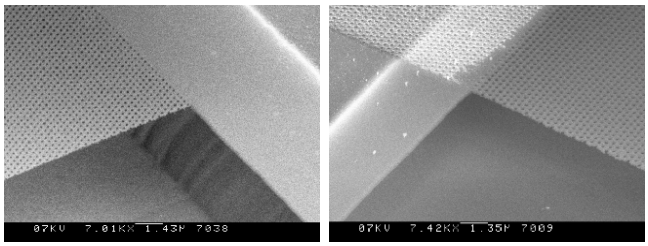


Figure 29. SEM micrograph of a high-porosity microsieve made with the lift-off method. Courtesy of Applied Optics Group, University of Twente.

with a product with a high affinity for the SAM species and a conformal SAM pattern is formed on the product (e.g., alkanethiol pattern on a gold coated product).PDMS is a rather elastic and relatively strong material very well suited for reproducible contacting purposes on nonplanar surfaces; however, it lacks a microporous microstructure for enabling functional fluid (ink) transport to the product domains or to enable other functional properties as will be described.

In Figure 30 three basically different printing techniques are represented.

Microporous stamps 21 (Fig. 30) with (alternating) regions with a dense skin layer and adjacent regions with a (porous) layer without the skin layer have easily been made with a phase separation process by locally removing the skin layer by, for example, oxygen plasma etching with the aid of a perforated mask shielding the remaining dense skin layer regions. The stamps are made with a phase separation process with the aid of a mold having patterned regions with sharp protrusions penetrating the microporous layers and patterned regions without such sharp protrusions where a dense skin layer is formed. In case the skin layer is not dense but nanoporous the skin layer can of course first be hermetically sealed without sealing the microporous part of the stamp with, for example, a hydrophilic coating (e.g., aliphatic and cyclic olefin-based polymers, or fluoropolymers or silicon based polymers). The stamps may also be subpatterned through use of photosensitive precursors in the casting solution of the product.

In one embodiment a (coplanar) stamp 11 with alternating nanoporous hydrophilic and dense hydrophilic surface regions is locally filled with an aqueous chromium etch solvent and brought into contact with a substrate having a chromium layer with a thickness of 20 nm. Whereas the dense regions 13 locally protect the chromium layer, in the nanoporous regions 12 an exchange between the chromium layer and the etch solvent may result in a locally dissolved and patterned chromium layer. Instead of chromium, many other materials or combinations of materials are applicable (e.g., aluminum, metal oxides and nitrides, metals, semiconductors, polymeric lacquer layers, etc.). The chromium layer may be replaced by a one phase lacquer layer and the solvent may be replaced by a second phase vulcanizing agent for the one phase lacquer layer. Instead of solvents also reactive gases can be used to etch patterns (e.g., SF_6 to etch and pattern silicon products). The microporous stamp can also be used to dab or adsorb locally a liquid or viscous layer that has been casted on a product. Dabbing may be improved by locally compressing the microporous regions during the contact of the stamp with the product. See Figure 31.

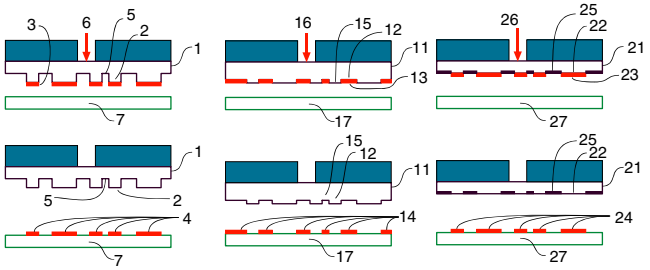


Figure 30. Left: the art of relief printing. The upper relief part 2 of the stamp 1 provided with a suitable ink coating 3 is contacted with a substrate 7 with a high affinity for the ink species and a conformal pattern 4 is formed on the substrate 7. The lower relief part 5 may be made ink repelling with a suitable coating (e.g., PVA, PVP) in order to avoid smearing of the pattern 4 of ink originating from sections 5. The upper relief part of the stamp 1 is provided with a macro- or nanoporous structure to contain ink or to transport ink from an injection point 6 for reproduction or continuous printing of the pattern 4 on the substrate 7. Middle: the art of gravure printing. The engraved part 12 of the stamp 11 provided with a suitable ink coating 13 is then contacted with a substrate 17 with a high affinity for the ink species and a conformal pattern 14 is formed on the substrate 17. The nonengraved part 15 may be made ink repelling with a suitable coating (e.g., PVA, PVP) in order to avoid smearing of the pattern 14 of ink from sections 15. The engraved part of the stamp 11 is provided with a macro- or nanoporous structure (cf. SEM picture below: engraved part is microporous, nonengraved part has a dense skin layer) to contain ink or to transport ink from an injection point 16 for reproduction or continous printing of the pattern 14 on the substrate 17. Right: the art of planographic printing (i.e., art lithography). The ink delivering part 22 and the nonink delivering part 25 of the stamp 21 are not determined by a difference in height but are made by the provision of suitable ink repelling and ink attracting coating s. The stamp 22 with a suitable ink coating 23 on part 22 is then contacted with a substrate 27 with a high affinity for the ink species and a conformal pattern 24 is formed on the substrate 27. Part 25 may be made ink repelling with a suitable coating (e.g., PVA, PVP) in order to avoid smearing of the pattern 24 of ink to sections 25. Part 22 is provided with a macro- or nanoporous structure to contain ink or to transport ink from an injection point 26 for reproduction or continuous printing of the pattern 24 on the substrate 27. Also in another embodiment part 25 may be microporous and filled with an ink repelling medium (e.g., water; Senefelder 1796).

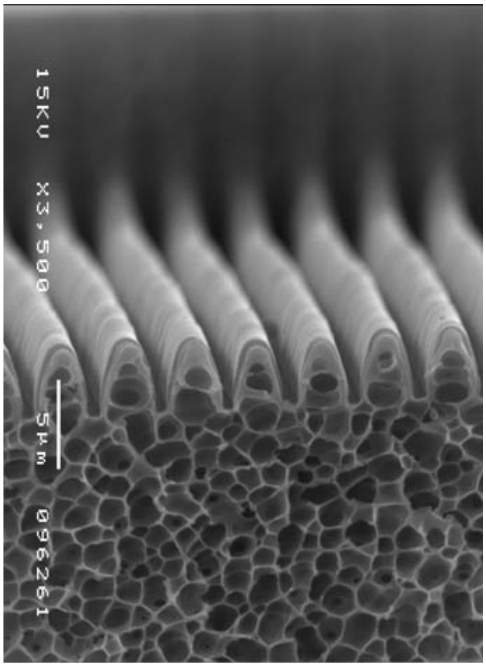


Figure 31. Left: Microporous stamp 1 produced with a phase separation process on a suitable mold. At left is a cross-section of a polyimide microporous microprinting tool with a smooth skin layer as obtained with a phase separation process. Courtesy of Aquamarijn Research/Membrane Technology Group, University of Twente.

In order to facilitate mask alignment of different mask stamping steps and to reduce thermal expansion differences between the stamps and the product, stamp regions parts are provided on a transparent (e.g., Pyrex, borosilicate glass) support material with the same thermal expansion coefficient as the substrate. Preferentially the microporous parts leading to the injection point have a high inner porosity to reduce flow resistance and a relatively small total dead volume in order to reduce the amount of adsorbed species.

In some cases it has proven to be useful to first print an ink pattern with the stamp on an intermediate dense or microporous transfer foil that transfers the ink pattern subsequently to the substrate, especially foils which have a well defined wetting contact angle with the selected ink medium.

The substrate may also first be provided with a suitable adsorbive nonsplattering or nanoporous (sacrificial) coating for more adsorption of the ink (e.g., obtained with a phase separation method). In another embodiment this nanoporous coating is made by deposition of an aluminum layer with a thickness of 200 nm on a silicon wafer and transforming this aluminum layer to a nanoporous (porosity 60–90%) honeycomb structure with thin vertical walls (pore spacings 10–50 nm) by anodic oxidation techniques well known in the art. Silicon and many other materials with different layer thicknesses can be transformed as well with anisotropic etching techniques for similar purpose. After the local deposition of the ink or an etch resistant lacquer in and/or on this layer, and preferentially dissolving the remaining uncovered layer, the substrate is ready for further processing steps.

Of course stamps may also be used for the formation of microstructures or microtransfer molding on planar and nonplanar surfaces of polymeric, ceramic, or metallic articles.

2.2. Fabrication of Micro/nanosieves with Silicon Wafers

Most of the processed silicon wafers today have a ⟨100⟩ orientation. The ⟨111⟩ planes etch approximately 200–1000 times slower than other oriented planes. This means that the ⟨111⟩ planes practically function as an etch stop. In order to obtain a silicon microsieve with a high degree of perforation, KOH etching must be carried out from the back side as well as through the pores on the front side. The etch rate on the front side decreases for smaller pores. For 1 μm pores it is roughly one order of magnitude smaller than the etch rate on the back side. If the wafer thickness is d, the width w of the openings on the back side must be approximately $d\sqrt{2}$ (depending on the etch rate underneath the membrane) to etch through the wafer, due to the 54.74° angle between the ⟨111⟩ planes and the wafer surface. For a 3-inch wafer with a 375 μm thickness this would result in membrane fields of approximately 530 μm width. Scaling up to 4-inch or 6-inch wafers with a thickness of 525 and 675 μm would give fields of 740 respectively 950 μm. As large fields are weaker than smaller fields, usually small additional support bars are etched underneath the membrane to reduce l (cf. Fig. 32).

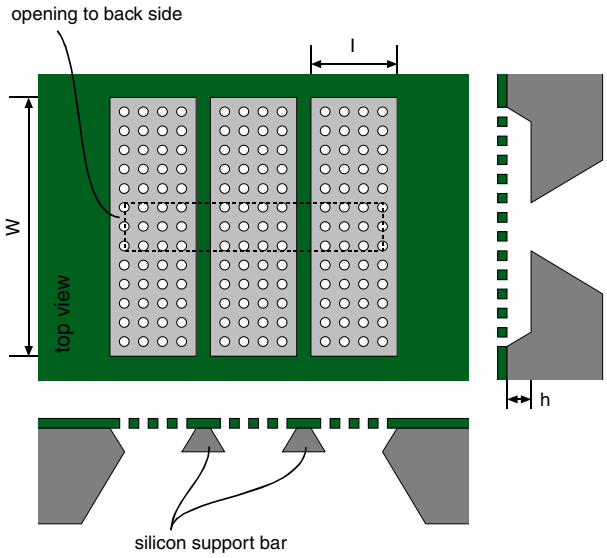


Figure 32. Top view of a membrane field that has been split up by small silicon support bars. The dashed line indicates the channel to the back side of the wafer.

2.2.1. Wet Etching through the Pores

However, for pore diameters below about 1 μm problems occur due to pressure buildup of the hydrogen gas that is created during KOH etching. Figure 32 shows that the silicon is etched through the pores. The gas can only escape through the pores in the hydrophilic membrane if its pressure exceeds the bubble-point pressure P_b, given by [80],

$$p_b = \frac{4\gamma \cos\theta}{d}$$

where γ is the surface tension of the KOH solution, θ is the liquid–solid contact angle, and d is the diameter of the pores. After the hydrogen gas has pushed the liquid out of a pore it will form a bubble on top of this pore. When the contact angle of the bubble with the pore wall becomes 0°, the gas pressure reaches a maximum value. For our KOH solution with an estimated value [81] for γ of 0.075 N/m, this reduces to

$$p_b = \frac{0.30}{d}$$

The bubble-point pressure increases for smaller pores. If the pore size is below a certain value, the gas can break the membrane. The maximum pressure $p_{\max}$ that an unperforated rectangular membrane can stand has been calculated by van Rijn et al. [115]:

$$p_{\max} = 0.69\frac{h\sigma_{\text{yield}}^{3/2}}{lE^{1/2}}$$

This equation shows that a decrease in size of the membrane field leads to a stronger membrane. A critical step in microsieve manufacturing is the release of the perforated membrane from the silicon support. Hydrogen gas that is created during KOH etching of the silicon builds up a pressure that might damage the membrane. Especially for submicrometer perforated membranes rupture is

likely to occur. Two different approaches were investigated to avoid the pressure problem. The first approach is based on plasma etching instead of KOH etching. Since no liquids are involved, the gaseous reaction products do not have to exceed a bubble-point pressure to escape through the pores. Using an SF_6/O_2 plasma and cryogenic substrate cooling, submicrometer perforated membranes were successfully released. A second approach is the formation of gas-escape channels to the back side of the wafer. This was achieved by using wafers with a ⟨110⟩ orientation, which allows for the possibility to etch channels with vertical sidewalls. With this second method also membranes with submicrometer pores were successfully released. The advantage over plasma etching is the possibility to process large batches. Furthermore, thanks to the vertical sidewalls thick (and thus strong) wafers can be used, while small (and thus strong) membrane fields are obtained. See Figure 33.

2.2.2. ⟨100⟩ Silicon Wafers

The base material of a microsieve is a ⟨100⟩ oriented silicon wafer. The wafer is coated with a silicon-rich nitride layer (intrinsic stress 10^8 Pa) by means of LPCVD. This layer is perforated using photolithography and reactive-ion etching with a CHF_3/O_2 plasma. Finally the silicon underneath the perforated layer is partially removed by anisotropic KOH etching to form a support.

This last step is crucial for the production of microsieves. If the silicon is etched from the back side with a KOH solution, a substantial part of the membrane will not be released due to the oblique ⟨111⟩ crystal planes in the ⟨100⟩ wafer. A much better result is obtained when etching takes place from the back side as well as through the pores on the front side. An additional advantage of this method is the possibility to etch extra support bars underneath the membrane (see Fig. 34).

The hydrogen bubbles that arise from KOH etching (25%, 70 °C) cause a pressure buildup underneath the membrane. This pressure may cause rupture of the membrane. An investigation was made of under what conditions KOH etching through the pores is still applicable and of the possibilities of dry etching to release the membrane.

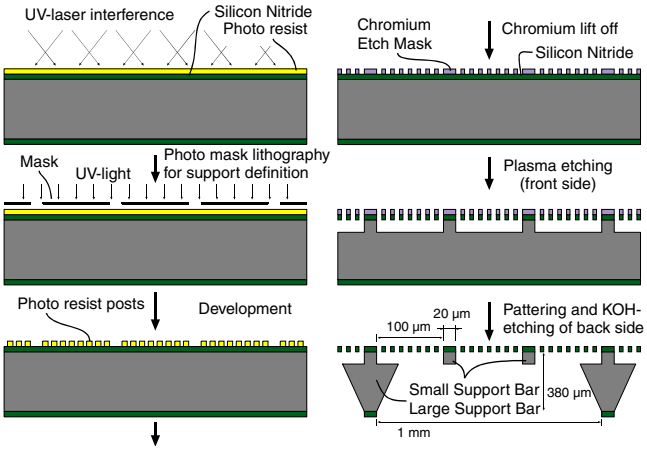


Figure 33. Fabrication process of a microsieve using laser interference lithography.

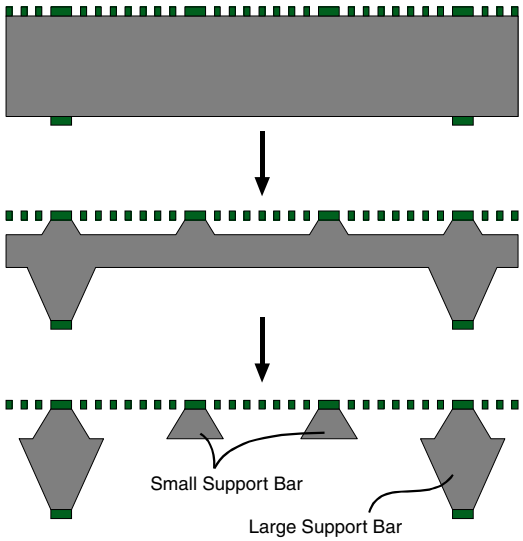


Figure 34. Releaseof the perforated membrane by anisotropic KOH etching from both sides of the wafer.

Dry Etching through the Pores In order to overcome the problems of wet etching, another possibility is releasing the membrane with the use of dry etching through the pores. The basic idea is given in Figure 35.

A patterned etch mask (photolacquer or chromium) is used to perforate the silicon nitride layer by CHF_3/O_2 etching. The mask is not removed from the nitride layer, as it will serve again as a mask for the silicon dry etching. The use of an isotropic etch gas is not satisfactory, because the small support bars would be etched away. Therefore an anisotropic-etch recipe is required, with just enough undercut to remove all the silicon between the pores. Plasma etching gives such an anisotropy that the ions can be accelerated

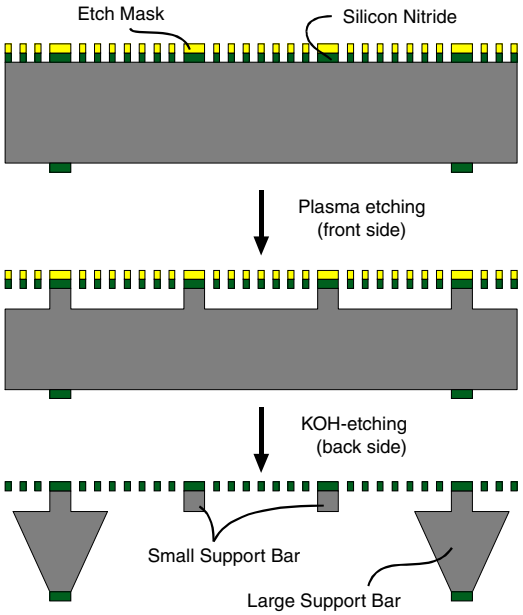


Figure 35. Process scheme for the release of perforated membranes by plasma etching.

into a vertical direction by an electric field. SF_6/O_2 mixtures are normally used, as SF_6 etches silicon isotropically while O_2 gives an anisotropic profile by passivating the silicon sidewalls of the trenches via LPCVD [83]. In order to obtain a higher etch selectivity between the silicon support and the silicon nitride membrane, the etch step was performed in an apparatus with cryogenic substrate cooling (Plasmalab 100, Oxford Plasma Technology [84–86]).

In contrast to KOH etching the depth of the channel underneath the membrane does not seem to be influenced by the pore size. The porosity will play a much more important role in the silicon etch rate, as it determines the amount of plasma that can enter the channel. See Figure 36.

2.2.3. ⟨110⟩ Silicon Wafers

The possibility of creating vertical channels in ⟨110⟩ silicon by KOH etching allows for the formation of small and therefore strong membrane fields. Buildup of a hydrogen pressure no longer occurs, as the gas can escape through the channels (see also Fig. 39). Furthermore the vertical anisotropy makes it possible to construct very thick and thus strong microsieves. For the new fabrication process a chromium etch mask is used instead of photolacquer. This chromium mask has an additional advantage of giving better defined pores, herewith improving the filter performance.

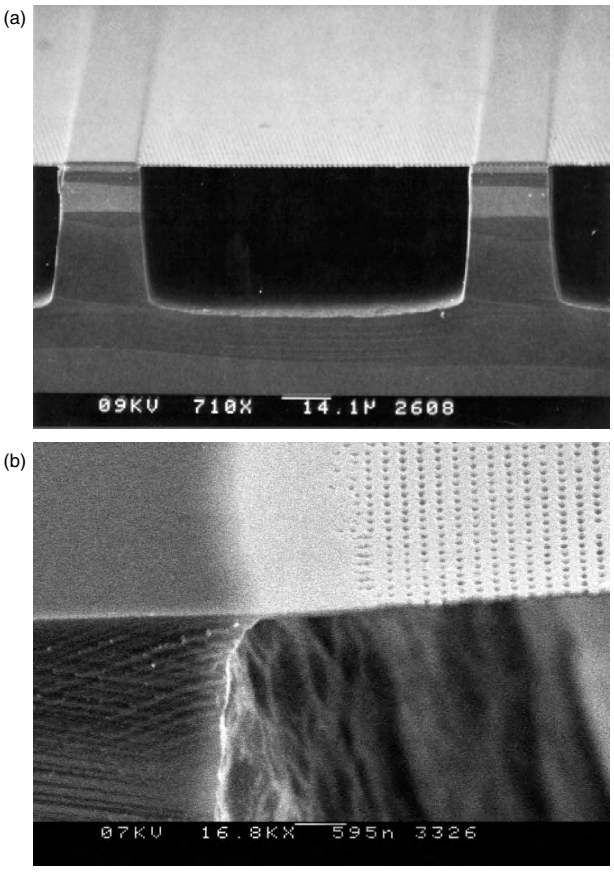


Figure 36. Shown are (a) 45 min, −130 °C, 400 nm pores and (b) 45 min, −130 °C, 70 nm pores, SF_6/O_2 etching through very small pores. Courtesy of Aquamarijn Research.

KOH etching on ⟨110⟩ wafers gives vertical walls but unfortunately two oblique ⟨111⟩ planes disturb the possibility to make small channels. However, as the shortest width of the fields determines the membrane strength, long but thin slit-shaped channels would still give strong membranes (see Fig. 37). For ⟨110⟩ wafers the walls of such channels have mutual angles of 70.53° and 109.47°. The walls are vertical, but in the sharpest corners (70.53°) the oblique planes arise with an angle of 35.26° relative to the horizontal ⟨110⟩ plane. A schematic view of such a channel is given in Figure 38. In this figure the silicon nitride on the front side has not been patterned yet, in order to etch the silicon only from the back side.

The channel length l has been chosen such that the two oblique planes intercept at the membrane layer on the top of the wafer before crossing each other. The front view in Figure 38 shows that the angle between the oblique planes and the wafer surface is 30° in the direction parallel to the long vertical walls. This means that in order to etch through the entire wafer the length of the channel has to be larger than $2d\sqrt{3}$. For the channel width w the only restriction that applies is given by flow-resistance requirements. The width should be chosen such that the channel resistance becomes much smaller than the resistance of the attached membrane field. It can be calculated that—even for extremely porous membranes—a width of 100 μm is sufficiently large.

Using the new ⟨110⟩ method the microsieves can be made much thicker—and thus stronger—than with the conventional ⟨100⟩ method, as the vertical anisotropy makes the membrane width independent of wafer thickness. Since the aspect ratio of the trenches can be as much as 600 it can be expected that etching of deeper grooves will not cause any problems for the KOH solution (25%, 70 °C) we used [87]. For concentrations above 30% the formation of ⟨311⟩ planes has been reported, which disturbs the possibility to etch deep trenches [88]. The maximum etch rate of the ⟨110⟩ planes occurs for a KOH concentration of 25% and is nearly twice the etch rate of the ⟨100⟩ planes [89].

As the slit-shaped channels can be made very small, they can be placed close together, thus creating membrane fields with small widths. Figure 39 gives a schematic illustration of the fabrication process of a microsieve on ⟨110⟩ silicon.

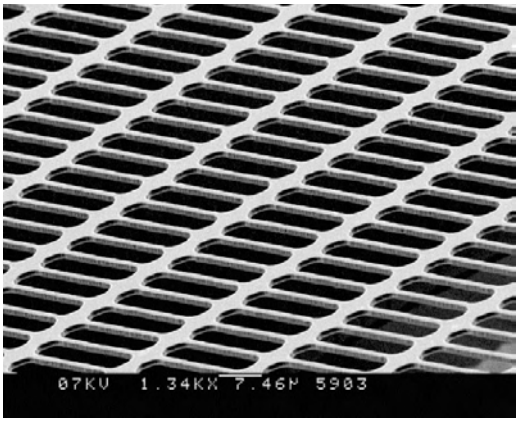


Figure 37. SEM micrograph of a microsieve with slit-shaped pores. Courtesy of Aquamarijn Research.

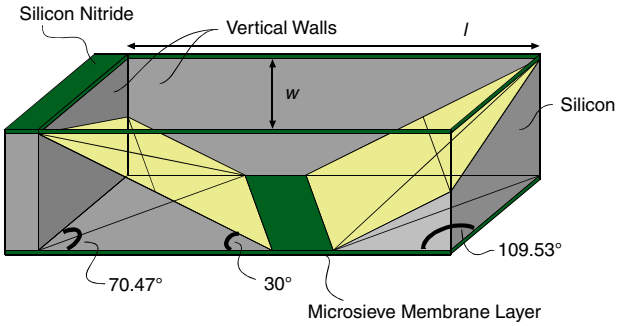


Figure 38. Schematic view of a slit-shaped pore after KOH etching of a ⟨110⟩ wafer. The shaded walls inside the channel on the right are ⟨111⟩ planes. Two planes are oblique and four planes are partly vertical.

Instead of a photolacquer mask a patterned chromium layer was used as a mask for plasma etching. This layer is obtained either by a lift-off process on photolacquer dots or by wet etching of the chromium through pores in a photolacquer layer. Before plasma etching of the membrane, vertical escape channels are etched in a KOH solution. Subsequently the membrane is patterned by plasma etching, using the chromium as the etch mask. Expected differences in etch rate due to variation in heat conduction between the released and unreleased parts of the membrane were not observed: the pores are uniform over the entire surface. For very small pores or very thin membranes it may be necessary to do an isotropic HF/HNO_3 etch to release the membrane and make a small space underneath it (third step in Fig. 39).

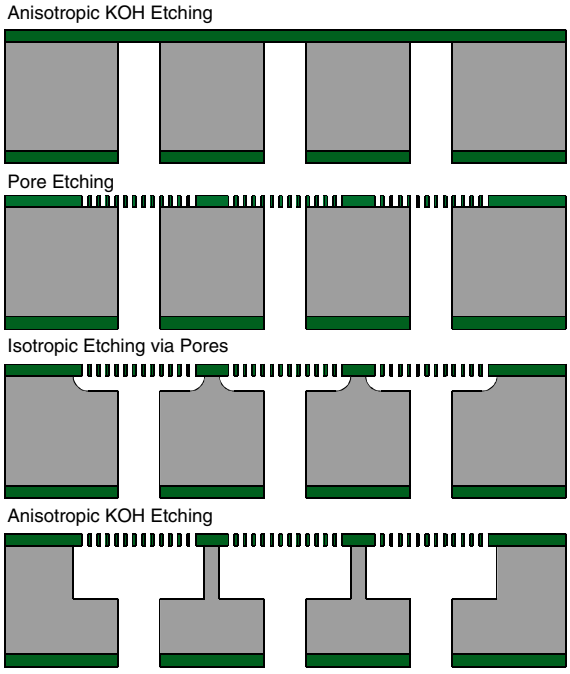


Figure 39. Schematic view of the microsieve fabrication on a ⟨110⟩ wafer. First vertical wall KOH etching in ⟨110⟩ silicon, then silicon nitride etching of the pores followed by release of the membrane with isotropic etching of silicon, and finally KOH etching to make the distance between the perforated membrane and the wafer larger.

If the HF/HNO_3 solution is chosen in the right composition (less then 1 part 50% HF on 100 parts 69% HNO_3), no hydrogen gas will be formed during this release process. The function of the small space underneath the membrane is to give the gas an escape route to the vertical channels. In the discussion we will calculate the required depth for this space.

Using the new method a microsieve was fabricated with slit-shaped perforations and a porosity of 75%. SEM micrographs of this sieve are shown in Figure 40. The shortest walls underneath the membrane should be partly vertical, but instead they are somewhat oblique with a rough surface. This is probably caused by the presence of the perforated membrane, as the back side of the wafer does show the vertical walls.

2.3. Research and Applications with Nanosieves

2.3.1. Patterning of Nanostructures

With laser interference lithography it is easily possible to reduce pore sizesa in microsieves to the nanometer regime, herewith giving birth to "nanosieves" [74]. A known technique to pattern surfaces on a substrate is to evaporate a material through a thin membrane (shadow mask) with well-defined openings [90] (see Figs. 42 and 43). It is also well known to use such shadow masks as a poor man's technique to make patterns on a substrate if photolacquer patterned layers are too elaborate or not wanted. Shadow masks with nanosized perforations may therefore be used in applications such as ion beam etching [91], electron beam patterning, near field optics, etc. Reactive ion etching through a shadow mask can be used for direct etching [3] of a (nano)pattern in a substrate (for example metal or polymeric foils).

Self-assembly (Nano)mask Preparation The micro/nanopattern may for instance be formed by using particles with a uniform size, for example a silica dispersion or a latex suspension, with particle sizes ranging from 5 nm to 5 μm.

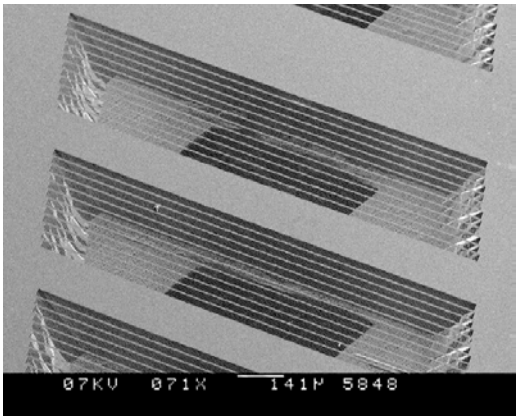


Figure 40. High-porosity microsieve fabricated on a ⟨110⟩ silicon wafer. Through the slit-shaped pores some typical etch angles can be recognized. Courtesy of Aquamarijn Research.

This suspensions may directly be formed on a substrate with use of a spin coating or evaporation technique. A more or less ordered distribution of particles will then be found in the pattern forming layer on the support. After the evaporation of the solvent of the suspension a very thin metal layer (e.g., a 10 nm chromium layer) may be deposited (e.g., by means of vapor deposition, i.e., sputtering or evaporation) on the substrate and on top of the particles, for example silica particles with a diameter of 30 nm. The silica particles are then solved in a buffered HF solution and a perforated chromium layer remains with perforations of approximately 20–30 nm in diameter depending on the chromium deposition conditions (correction for shadow effect of deposition). The chromium layer may be used as a membrane layer or alternatively as a mask layer for the (dry) etching of a membrane layer underneath the chromium layer (see Fig. 41).

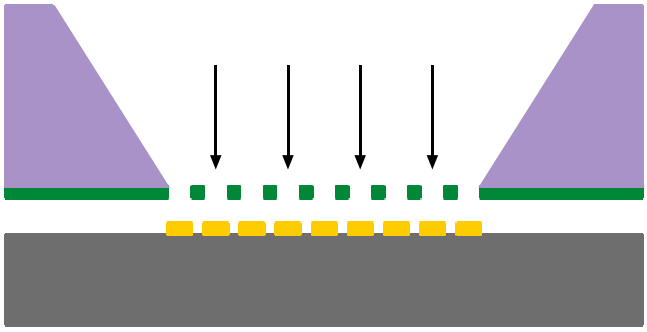


Figure 42. A nanosieve may be used as a shadow mask (nanostencil) for evaporation of material (e.g. gold) through the pores of the nanosieve. Reprinted with permission from [92], J. Brugger et al., *Microelectron. Eng.* 53, 403 (2000). © 2000, Elsevier Science.

2.3.2. Nanosieves for Photonic Structures

Recently, there has been growing interest in photonic bandgap materials, which have a spatially periodic refractive index with a lattice constant on the order of the wavelength of light, because of potential scientific and technological applications based on their unique light propagation properties [94]. However, experimental surveys of the optical properties of photonic bandgap materials have been limited due to the difficulty in preparing well-constructed samples except in the long wavelength region [95, 96].

Photonic crystals are novel materials with unique optical properties [97]. The crystals have a periodic modulation of the refractive index. As a result, the dispersion of light will be described by a band structure analogous to those of electron wave atomic crystals. Under the right conditions a photonic crystal can exhibit a photonic bandgap: light in a certain range of optical frequencies is forbidden in the crystal [98]. The existence of a photonic bandgap enables an unprecedented control of spontaneous emission and propagation. By locally disturbing the periodicity, a defect-associated photon state is created which can be used to guide light.

A photonic crystal slab is a thin film with a two-dimensionally periodic refractive index modulation in the plane [100, 101]. In a photonic crystal slab the light is confined to the crystal plane by a classical slab waveguide construction. For in-plane wave vectors a bandgap can be created. Thus, the slab has applications in light guiding without the need for a full bandgap in all three dimensions. Optimal performance of the photonic crystal slab is expected when the slab is mirror symmetric in the vertical direction (i.e., when the material on both sides of the slab has equal refractive index at least in the region of the near-field tail of the in-plane propagating light) [102].

Photonic crystals are fabricated by periodically arranging materials with highly dissimilar refractive index. To obtain a bandgap in the visible the periodicity of the index modulation has to be in the submicrometer range, <350 nm. For the fabrication of photonic crystal slabs electron beam lithography has been by far the most frequently used technique, since it is one of the few techniques that fulfills the accuracy requirements on these length scales [103]. However, due to its "direct sequential write" character e-beam lithography is time consuming for large area structures, plus uniformity errors in periodicity can occur due to errors (or quantization roundings) in the beam alignment. It would be highly advantageous to define the entire periodic pattern at once. Laser interference lithography (LIL) has proven its ability to generate uniform periodic submicrometer patterns over very large areas (on the order of square centimeters) [74, 104]. A combination of large area patterning with LIL and local patterning with a direct writing method (e.g., a focused ion beam [105]) is ideal for future fabrication of photonic crystal slabs with arbitrarily shaped light-guiding defects.

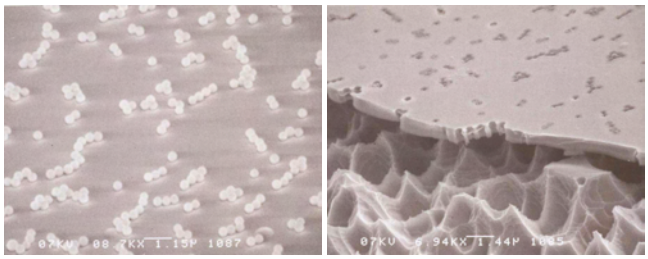


Figure 41. Photographs of a membrane made with a self-assembled mask. Left: arrays of nanosized particles on a silicon nitride membrane layer. Right: after chromium deposition over the particles and removal of the particles, the pores in the membrane have been etched. Reprinted with permission from [3], C. J. M. van Rijn, membrane Filter as well as a Method of Manufacturing the Same, PCT Application 95/1386206.

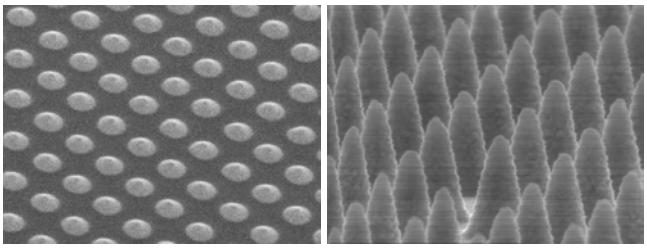


Figure 43. SEM micrographs of gold spots and pillars evaporated through a nanostencil. Reprinted with permission from [93], M. Kölbel et al., *Nanoletters* 2, 1339 (2002). © 2002, American Chemical Society.

In Figure 44 a combination of LIL with focused ion beam (FIB)-assisted deposition is presented. Highly uniform freestanding photonic crystal slabs have been fabricated, extending over areas as large as a few hundred micrometers squared, with the possibility to introduce any kind of defect structure. In this way is realized, to our knowledge, the largest photonic crystal slab (100 μm $\times$ 4 mm) for visible radiation. Long line defects, extending over more than 1 mm, were also introduced.

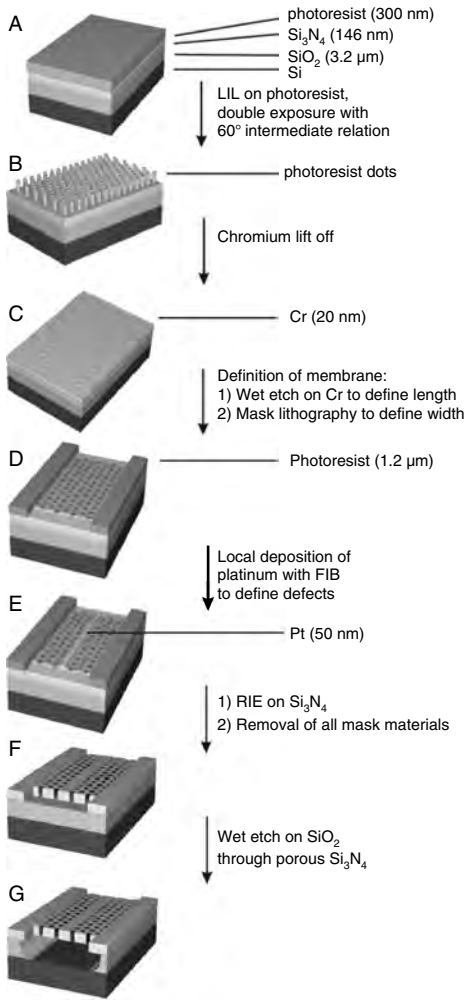


Figure 44. Schematic representation of the procedure which combines large area patterning through laser interference lithography and local introduction of defect structures through FIB-assisted deposition. The dimensions in these pictures are chosen for illustrative clarity and are not to scale. Note that a slab of 100 μm $\times$ 4 mm contains over 4.5 million holes. A photoresist layer on top of a Si_3N_4 (146 nm), SiO_2 (3.2 μm), and Si layer stack (A) is converted into a large area triangular dot pattern with LIL (B). By chromium lift-off a thin, highly selective etch mask containing holes is formed (C). Next, the etch mask is divided into smaller sections by mask lithography (D). Local deposition of platinum is used to define defect structures of any shape in the photonic crystal slabs. Here, we demonstrate the introduction of a line defect (E). By RIE the structure is etched into the Si_3N_4 layer, after which Pt, Cr, and photoresist are removed (F). By a wet etch through the holes the underlying SiO_2 layer is removed and the freestanding Si_3N_4 photonic crystal slabs are formed (G). Reprinted with permission from [99], L. Vogelaar et al., *Adv. Mater.* 13, 1551 (2001). © 2001, Wiley-VCH.

In the photonic crystal slabs Si_3N_4 ($n = 2.16$ at $\lambda =$ 670 nm) and air acted as high and low index material, respectively. A triangular lattice of air holes in the Si_3N_4 layer was fabricated in order to obtain a bandgap for TE modes (polarization in the plane of the crystal). By making the slab freestanding (i.e., embedded in air on both sides), we obtain the desired vertical mirror symmetry of the crystal slab. See Figures 45 and 46.

The advantage of LIL is that the entire periodic pattern is created at once. A high accuracy and flexibility of FIB deposition to introduce line defects in the extended crystals has been used. In principle, any defect is possible. The combination of LIL with FIB deposition opens avenues for more optical applications of photonic crystals. Due to its simplicity and scope, FIB-assisted LIL is a promising alternative for e-beam lithography in the fabrication of photonic materials. See Figure 47.

***Alumina Photonic Crystals* [*106*]** Anodic porous alumina with an ideally ordered air-hole array with a 200 to 250 nm lattice constant exhibits a two-dimensional (2D) photonic bandgap in the visible wavelength region. The air-hole preparation process was based on Al pretexturing by imprinting with a mold (cf. Fig. 66 in Section 4.3), followed by the anodization of Al in acid solution.

The depth limit of the ordered air holes is due to the disturbance of the arrangement of the air holes along with the growth of the oxide layer during anodization. For the application of 2D photonic crystals to a wide range of scientific and technological fields, an air-hole array with a higher aspect ratio is required. Anodization in a phosphoric acid solution under a constant voltage condition of 195 V yielded an air-hole configuration with a lattice constant of 500 nm. See Figures 48 and 49.

Figure 50 shows the typical transmission spectra of the anodic porous alumina with a 500 nm lattice constant. The filling factor of the air holes was 0.1. In Figure 50a, which presents the transmission spectra for H polarization light, distinct dips in the transmission were observed for both the J and X directions. The observed position and width were in good agreement with the theoretical prediction based on the plane-wave method for the triangular lattice of air holes in the alumina matrix [107].

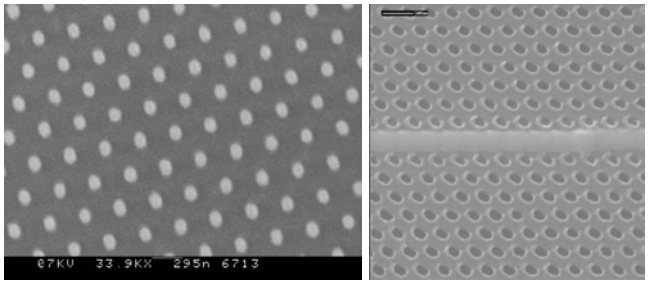


Figure 45. Left: SEM image of triangular photoresist dot pattern with a period of 300 nm after laser interference lithography. The uniform pattern extends over an area of the order of a centimeter squared. Right: FIB image of an approximately 50 nm thick platinum line covering holes in a 20 nm thin chromium mask. The platinum line defines a defect structure in the photonic crystal slab and is deposited locally with a focused ion beam. Courtesy of Adv. Materials.

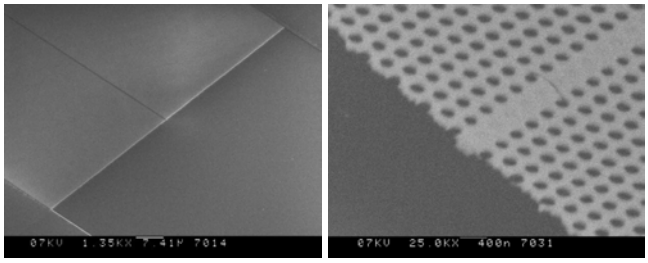


Figure 46. Left: SEM micrograph of the end face of a photonic crystal slab between the support bars. This 146 nm thin slab is freestanding across 100 μm and has a length of 4 mm. Right: SEM micrograph, close-up of the defect line in the photonic crystal slab. The highly uniform triangular hole pattern has a period of 297 nm. The axial widths of the elliptical holes are 212 and 164 nm. The shape of the holes appears unharmed by the presence of the defect line. The defect is expected to form a multimode waveguide for visible light (estimated central wavelength 670 nm). Right: Courtesy of Adv. Materials. Left: Reprinted with permission from [99], L. Vogelaar et al., *Adv. Mater.* 13, 1551 (2001). © 2001, Wiley-VCH.

2.3.3. Nanoengineered Membrane Separation Technology

Introduction Nanosieves may be used as a stencil for patterning or as a photonic crystal, but they may also be used as a proper solid support for much thinner membrane layers, or to put it simply: as a thin functional membrane in the pores of another supporting membrane structure.

The advantage of this nanoarchitecture is that a decoupling is made between the thickness of the functional membrane layer and the strength of the support membrane layer. This nanoarchitecture will be of use in processes (e.g., gas separation, ultrafiltration, etc.) in which the flux through the functional membrane layer is diffusion limited. It even would be possible to obtain gas separation/molecular sieve membrane layers (e.g., silicon oxide, silicalite) with a thickness of a few nanometers that would enable gas molecules

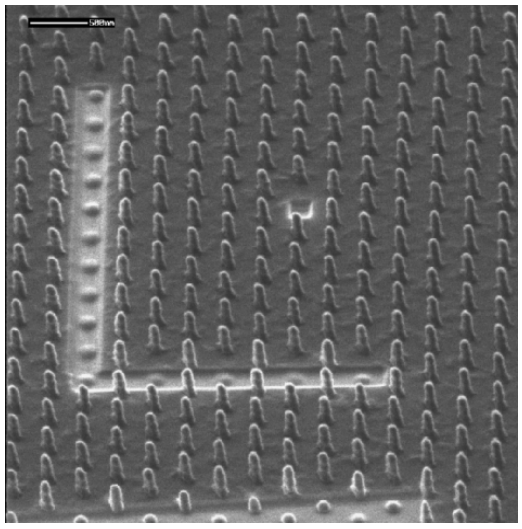


Figure 47. SEM micrograph of a FIB-assisted removal of a number of pillar structures obtained with laser interference lithography. Courtesy of Applied Optics Group, University of Twente.

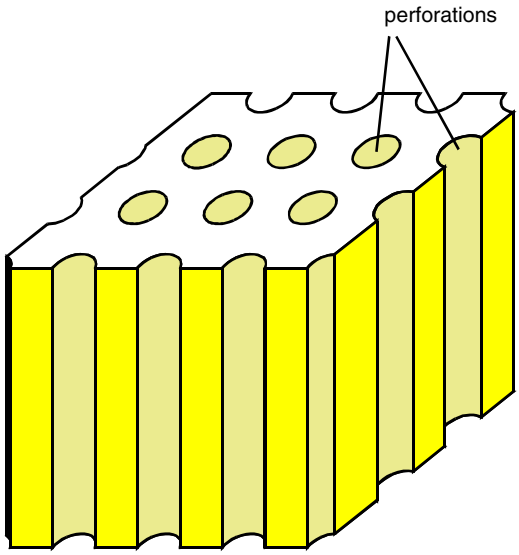


Figure 48. Schematic diagram of a 2D photonic crystal using anodic porous alumina.

to pass the membrane ballistically [108] without colliding with the walls inside the membrane. Very high fluxes may be expected in this case.

An estimate of the maximum pressure load q of such a thin functional membrane with thickness h, width l, Youngs modulus E and yield stress σ may be obtained with an

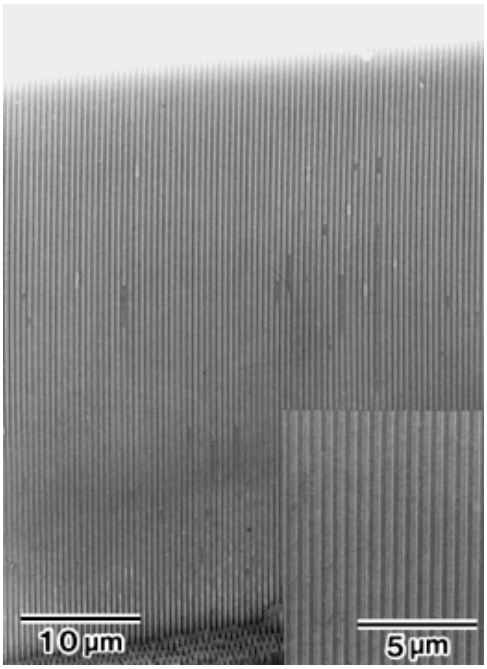


Figure 49. SEM micrograph of the highly ordered anodic porous alumina. Anodization was conducted in 0.1 M phosphoric acid solution at 0 °C at 200 V for 360 min. The lattice constant and filling factor were 500 nm and 0.3 μm respectively. The inset shows an enlarged view of the air-hole array at the bottom region of the anodic porous alumina. Reprinted with permission from [106], H. Masuda et al., *Jpn. J. Appl. Phys.* 39, 1039 (2000). © 2000, The Japan Society of Applied Physics.

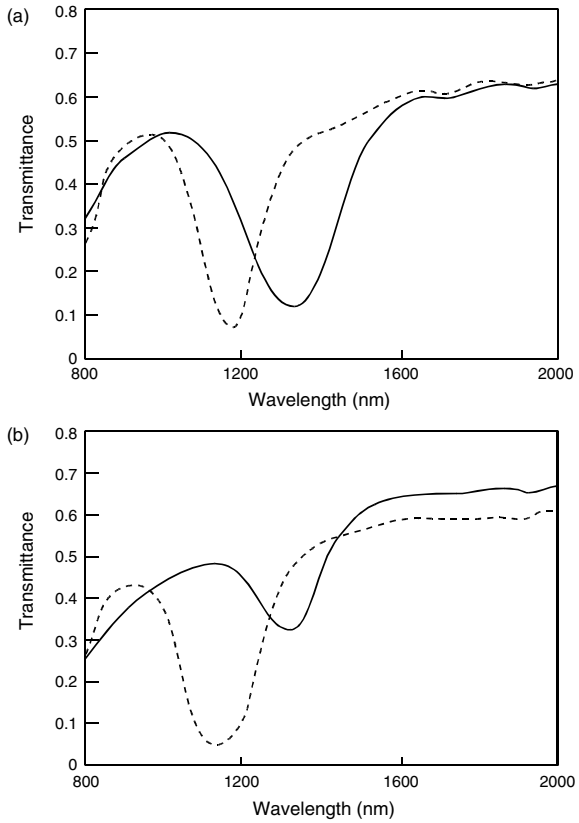


Figure 50. Transmission spectra of anodic porous alumina for (a) H and (b) E polarization light for the Γ–J (- - -) and Γ—X (—) directions. The lattice constant and filling factor were 500 nm and 0.1 respectively. Reprinted with permission from [106], H. Masuda et al., *Jpn. J. Appl. Phys.* 39, 1039 (2000). © 2000,

earlier expression:

$$\sigma_{\text{total}} = \sigma_{\text{tensile}} + \sigma_{\text{bend}} \underset{x=0}{\overset{u \gg 1}{=}} 0.29(1 + 1.47/0.37)\sqrt[3]{\frac{q^2 l^2 E}{h^2}}$$

Using current values for a silicon oxide membrane layer with a thickness of only 1 nm and width of 1000 nm, a maximum pressure load is obtained of 16 bar, a value comparable with (e.g., a silicon nitride layer with a thickness of 1 μm and width of 200 μm) the maximum pressure load of the membrane support layer itself. See Figure 51.

Microengineered membranes have also been the subject of study at other institutes for example, for air filtration and particle counting applications [109]. Following earlier work of van Rijn [3] palladium membranes for hydrogen separation have also been recently made using a microsieve structure as a support for a very thin palladium membrane (<200 nm) for, for example, small scale fuell cell applications [110]. Also membranes with straight nanometer sized pores (20 nm) have recently been made for cell immunoisolation and biomolecular separation applications [111–113] following formerly presented nanopore construction technologies [3] with use of conventional lithographic techniques (see Fig. 52).

In a membrane layer 2 of silicon nitride with thickness 1 μm parallel grooves with a depth of 1 μm and a length of

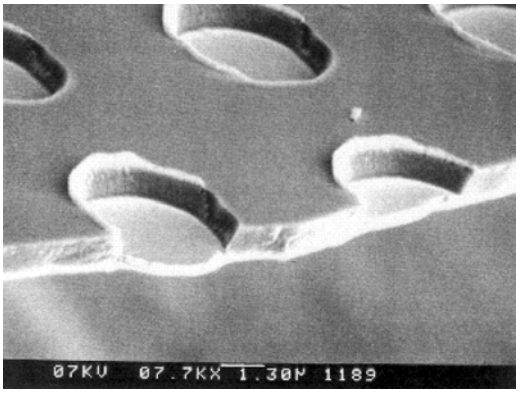


Figure 51. Pd membranes with a thickness of 200 nm and diameter 4 μm hanging in the pores of a silicon nitride membrane with a thickness of 1 μm. Reprinted with permission from [3], C. J. M. van Rijn, Membrane Filter as well as a Method of Manufacturing the Same, PCT Application 95/1386206.

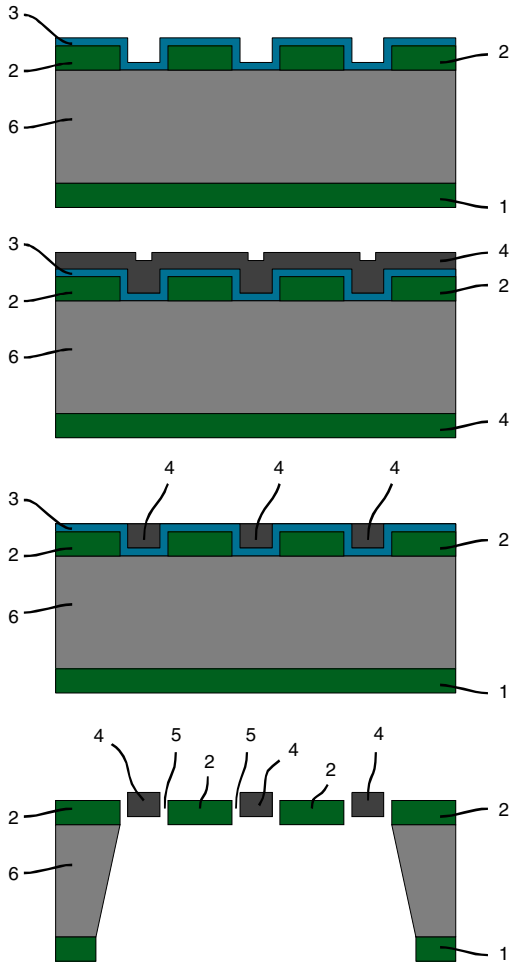


Figure 52. In cross section, subsequent stages of a method of manufacturing a membrane filter with nanosized pores. Reprinted with permission from [3], C. J. M. van Rijn, Membrane Filter as well as a Method of Manufacturing the Same, PCT Application 95/1386206.

10 μm are being etched with reactive ion etching techniques. A sacrificial layer 3 of silicon dioxide is grown uniformly with a uniform thickness of 20 nm by means of chemical vapor deposition of tetraethoxysilane at elevated temperature. Square to the grooves a line pattern (not shown in the figures) is etched in the sacrificial layer 3 with a depth of 20 nm, a width of 1 μm, a length of 1000 μm, and a spacing of 10 μm with standard lithographic techniques using a buffered HF solution as a selective etchant. Over the sacrificial layer 3 and in the grooves a groove filling layer 4 with a thickness of 2 μm is deposited (e.g., a polysilicon layer). The groove filling layer 4 is then planarized by polishing or etching such that the grooves remain at least partially filled with groove filling material 4. The groove filling material 4 is then at least partially directly connected to the membrane layer 2 near the line pattern. Subsequently, the sacrificial layer 3 is removed from using the buffered HF solution such that channel-like pores 5 with a width of 20 nm, a depth of 1 mm, and a lateral length of 10 μm are realized. Openings are etched in the support 6 using KOH as an etchant and the patterned silicon nitride layer 1 as an etch mask. With the process it is possible to make very small pores without alignment of the mask for the parallel grooves and the mask for the line pattern.

An intermediate layer between the membrane layer 2 and the support 6 may be applied as a selective etch stop, for example silicon nitride or alumina.

The membrane may also be used as a separation filter (potential low diffusion times), for instance for *in vivo* screening of blood cells, or in a sensor [3, 114] or actuator system. This will give a clear advantage for microsensors and actuators that are processed with use of thin film technology.

Desai et al. [113] have studied the use of nanoporous membrane filters with pore sizesa of 25 and 50 nm for biosensor applications according to the manufacturing scheme of Table 5. Some typical diffusivity parameters found with their membranes in comparison with conventional membranes are listed in Table 6.

2.3.4. Microfiltration of Lager Beer

The Grolsche Bierbrouwerij Nederland and Aquamarijn Micro Filtration started a collaboration in 1995 with the purpose of investigating the possibilities of replacing diatomaceous earth (kieselguhr) filtration of lager beer with microsieve filtration [115].

Nowadays tubular ceramic or polymeric membrane systems are being developed as an alternative for classical packed bed diatomaceous earth (see Fig. 1) dead-end filtration. Complications in using these type of filtration membranes are yeast cell clogging and protein adsorption leading to a fast flux decline and subsequent elaborate in-line cleaning procedures. A pilot plant was built in which the performance of the microsieves was tested. Clarification of lager beer is an important operation during the brewing process. Rough beer is filtered in order to eliminate yeast cells and colloidal particles responsible for haze. Common beer-filtration systems are based on kieselguhr. However, the exploitation costs of these systems are rather high. See Figure 53.

Cross-flow microfiltration with polymeric or ceramic membranes may be an alternative. Several studies have been carried out, but often problems like poor permeate quality (i.e., high turbidities or protein and aroma retention) or insufficient fluxes are encountered [116, 117]. Experiments at the Grolsch breweries in which a microsieve was used for the filtration of beer [118] showed a permeate flux of 4×10^3 l/m^2/hr during a period of at least 5 hours without any increase in transmembrane pressure (see Fig. 54).

This flux is one to two orders of magnitude higher than typical fluxes obtained with diatomaceous earth or other membranes. In the experiments the formation of a cake layer was diminished by using a cross-flow configuration in combination with backpulse techniques and a transmembrane pressure of only 20 cm H_2O (0.02 bar). As the microsieve is made of an inert material it may be cleaned with aggressive chemicals or by steam sterilization. After cleaning the new permeate still has the tendency to foam. Many other membranes may give problems on this point, as it is difficult to remove cleaning agents from the large inner surface and from dead-end pores. An additional advantage is the absoluteness of the filtration: the uniform pores do not permit a single yeast cell to pass through the sieve. See Figure 55.

Moreover, extensive cleaning procedures are required, as beer turns out to cause severe fouling [117, 119, 120]. Ceramic membranes have an advantage over polymeric membranes regarding fouling, as they can withstand harsh cleaning methods. However, the obtained fluxes are usually significantly lower. Ceramic membranes with a small flow resistance would therefore be highly desirable for beer filtration. Microsieves made with silicon micromachining consist of a thin microperforated silicon nitride membrane attached to a macroperforated silicon support. The membrane thickness is of the order of the pore size, thus allowing high fluxes and relatively simple cleaning procedures. Moreover, the membrane is optically flat and smooth (surface roughness typically below 10 nm), which hampers adsorption of foulants. Furthermore, the pores are uniform in size and distribution, which may be important for quality control.

Table 5. Interaction of albumin with membranes is also measured by looking at the difference between the initial albumin concentration and the final concentration; the differences can be attributed to both protein diffusion and adsorption [113].

	Whatman (albumin concentration)		Millipore (albumin concentration)		Micromachined (albumin concentration)	
Time (min)	Absolute	(g/dl)	Absolute	(g/dl)	Absolute	(g/dl)
0	0.381 ± 0.004	3.88 ± 0.02	0.423 ± 0.003	4.31 ± 0.03	0.395	3.980 ± 0.003
420	0.002	3.58 ± 0.02	0.398 ± 0.003	4.05 ± 0.03	0.394	3.970 ± 0.002
Change in albumin concentration	0.352 ± 0.004	0.30		0.26		0.01

Table 6. Membrane parameters and calculated diffusivities [113].

	Whatman	Millipore	Micromachined
membrane thickness (μm)	1 (asymmetric pores)	105	5
Porosity (%)	50	70	0.86
Effective area (mm^2)	4	5.6	0.041
Effective diffusivity	4.00E−07	5.292E−05	1.025E−08
Absolute diffusivity ($D_{\text{eff}}/A_{\text{eff}}$)	1.00E−07	9.45E−06	2.50E−07

Membrane fouling during filtration of lager beer with microsieves can well be studied through in-line microscopic observations. Ph.D. student S. Kuiper [121] was willingly asked in 2000 by Aquamarijn to perform a number of experiments with lager beer from The Grolsche Bierbrouwerij at Enschede. All microsieves with a pore size less than 1.2 μm for this study were manufactured by Aquamarijn (with thanks to Dimes at Delft for the photolithography).

It was observed that the main fouling was caused by micrometer-sized particles, presumably aggregated proteins and/or polysaccharide/cellulose residues. These particles formed flocks covering parts of the membrane surface. Most of the flocks could be removed by a strong temporary increase in cross-flow. Underneath the flocks a permanent fouling layer was formed inside the pores. This made frequent removal of the flocks crucial in delaying the process of permanent in-pore fouling.

Besides the fouling process the influence of pore size on permeate flux and turbidity was investigated. Centrifuged beer appeared to give a significantly clearer permeate than rough beer. For centrifuged beer and a microsieve with a pore diameter of 0.55 μm a haze of 0.23 European brewery convention (EBC) was obtained during 10.5 hours of filtration at an average flux of 2.21×10^3 l/m^2 hr. For a sieve with slit-shaped perforations of 0.70×3.0 μm^2 a haze of 0.46 EBC was obtained during 9 hours of filtration at an average flux of 1.43×10^4 l/m^2 hr. This flux is about two orders of magnitude higher than is commonly obtained with membrane filtration of lager beer. Concentration of the beer by a factor of 12 in a 3 hour run hardly influenced the magnitude of the flux. See Table 7.

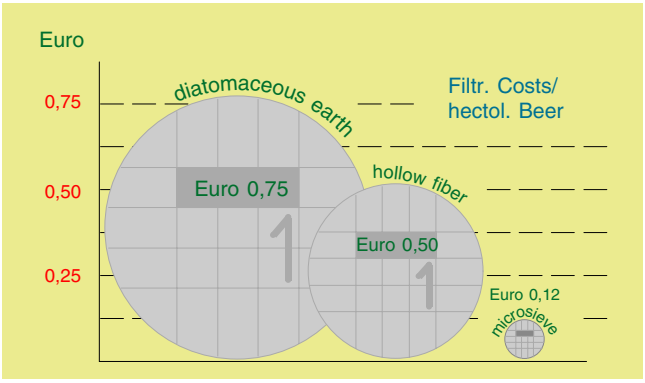


Figure 53. Global cost comparison of yeast cell filtration methods. Reprinted with permission from [115], C. J. M. van Rijn et al., *Proc. EBC Congress* 501 (1997). © 1997,

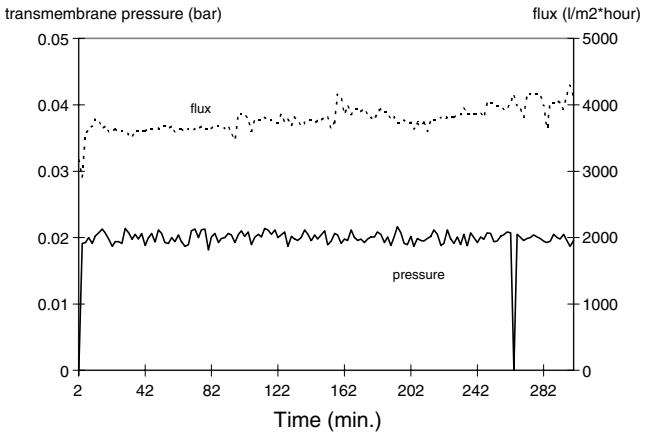


Figure 54. Behavior of flux and pressure in yeast-cell filtration of lager beer with a microsieve. Reprinted with permission from [115], C. J. M. van Rijn et al., *Proc. EBC Congress* 501 (1997). © 1997,

3. NANOENGINEERED BIOMEMBRANES

3.1. Introduction

The research goal in biomembrane science is to study the function of natural biomembranes and to engineer new types of biomembrane materials on the nanometer scale [122–124]. Natural biomembranes, in particular lipid bilayer cell membranes, have an important function in biological molecular exchange (metabolism) and signal transduction processes (e.g., immune reactions, hormone detection) of the living cell. Generally, these processes are mediated across the cellular membrane wall by a transmembrane channel and pores for molecular exchange [125–127] and for signal transduction [128, 129].

Most gram-negative bacteria possess an inner and an outer cellular membrane wall. The inner one contains a

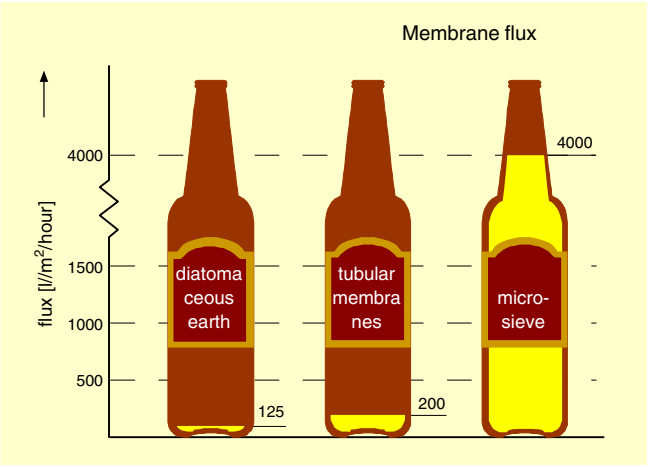


Figure 55. Flux comparison of yeast cell filtration methods. Reprinted with permission from [115], C. J. M. van Rijn et al., *Proc. EBC Congress* 501 (1997). © 1997,

Table 7. Results of three long-run experiments.

	Porosity (%)	Average beer flux (l/m^2 hr)	Initial beer flux (l/m^2 bar hr)	Water flux at 20 °C (l/m^2 bar hr)	Permeate haze (EBC)
0.70 μm slits	31	$14.3 \cdot 10^3$	$18 \cdot 10^4$	$18 \cdot 10^5$	0.46
0.80 μm circles	22	$7.24 \cdot 10^3$	$5.6 \cdot 10^4$	$5.8 \cdot 10^5$	0.58
0.55μm circles	24	$2.21 \cdot 10^3$	$1.8 \cdot 10^4$	$4.7 \cdot 10^5$	0.23

"dense" lipid bilayer, which is electrically closed, so that no ions or other hydrophilic substrates can cross the barrier without the help of highly specific membrane proteins. This inner cell wall is separated from the outer one by an aqueous phase filled with water-soluble polymers, the peptidoglycan. In contrast, the outer cell wall is fairly permeable to smaller solutes below a molecular weight of about 400 Da. Such substances can freely permeate under a concentration gradient through general diffusion porins in the outer cell wall. The most prominent of the general diffusion porins is OmpF (outer membrane protein F) [130, 131] which is very stable and does not denaturate, for example, in 4 M GuaHCL, 70 °C or in 2% SDS. In case of lack of nutrition, the pure diffusion process is too slow and the bacteria need to enhance the efficiency of the translocation functions. For those cases, nature has created a series of rather specific and highly sophisticated membrane channels. An extensively studied example is the malto-oligosaccharide specific channel LamB, or maltoporin [132] of *Escherichia coli*. Nanoengineered porins in artificial lipid or also block copolymer biomembranes may allow observation of the crossing of a single substrate molecule through the porin channel. A well studied example [133] is the α-hemolysin porin in a diphytanoylphosphatidylcholine (lipid) bilayer. See Figure 56.

High-resolution measurements of ion currents through single porin channels to probe, for example, neutral solute transport [134] have already been shown. Also the screening of facilitated transport of a wide variety of molecules [135], for example, to study antibiotic uptake through porins, to test for toxins, etc., is an object of ongoing interest.

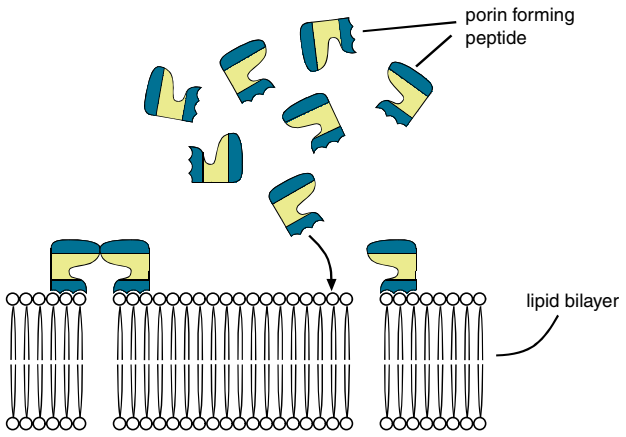


Figure 56. Schematic illustration of the assembly of a pore channel with (α-hemolysin) peptide oligomers in a (diphy-tanoylphosphatidylcholine) lipid bilayer. Intermediate structures of the assembly process are not considered.

3.1.1. Surface Layers

In microorganisms of almost every taxonomic group of walled eubacteria and archaeobacteria [136, 137] the outer cell wall has a nanoporous crystalline surface layer (S-layer). This surface layer is composed of a single protein or glycoprotein species with molecular weights ranging from 40,000 to 200,000 and can exhibit either oblique, square, or hexagonal lattice symmetry [138–140]. See Figure 57.

An important feature of these nanoporous structures is their repetitive physicochemical properties. Pores passing through S-layers show identical size and morphology and are in the range of ultrafiltration membranes. Functional groups on the surface and attached in the pores may be well aligned and are accessible for specific molecules in a very precise fashion. S-layer proteins also reveal the ability to self-assemble into two-dimensional crystalline arrays in suspension, on solid supports, at the air/water interface, and on lipid films [141].

Modification of S-layers [145] would allow several possibilities including the manipulation of pore permeation properties, the introduction of switches to open and close the pores, and the covalent attachment to surfaces or other macromolecules through defined sites on the S-layer protein. S-layer proteins can also be produced in large amounts by continuous cultivation of S-layer-carrying organisms. S-layer proteins can be considered biopolymeric membranes with properties ideally tailored by nature for many biotechnological applications [146], as matrices for the immobilization of a variety of materials including biologically active macromolecules [147, 148] and for functionalization of synthetic materials. The regular pore properties of S-layers are in contrast with the amorphous pore structure of phase inversion membranes and may therefore be used to produce artificial biomembranes showing excellent molecular mass cutoff separation capabilities [149]. See Figure 58.

One of the most common organisms used for producing S-layer-carrying cell wall fragments and S-layer protein for biotechnological applications is *Bacillus stearothermophilus*

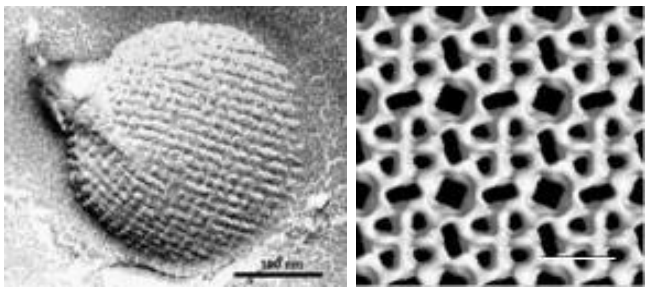


Figure 57. SEM photograph of a liposome with a nanoporous cell envelope (surface layer). Right: a reconstruction of an S-layer structure (bar = 10 nm). S-layers are formed by self-assembly and are made of protein and glycoproteins (MW 40,000–200,000) with an oblique, square, or hexagonal 2D crystal structure with typical lattice constants between 5 and 30 nm and a thickness of 5–10 nm. Courtesy of Nanosearch membrane GmbH.

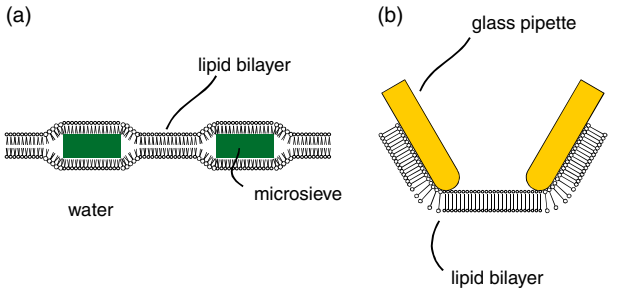


Figure 58. (a) Schematic illustration of a "black" lipid membrane generated by the method of Montal and Muëller on a micro/nanosieve. Originally this method involved the generation of a lipid bilayer over a septum with an orifice of 40–800 micrometers [142, 143]. For this purpose either a small drop of lipid dissolve in alkane is placed on the opening of the septum [144] or the membrane is formed from two lipid monolayers at an air/water interface by the position of the hydrocarbon chains through an aperture made in a hydrophilic partition which separates the two monolayers. (b) Schematic illustration of a lipid bilayer, generated on the tip of a patch clamp pipette.

PV72, for which a synthetic growth medium has been developed [150, 151]. S-layers have also been used for the production of S-layer based ultrafiltration membranes (SUMs) [152]. SUMs are produced by S-layer self-assembly or by depositing S-layer-carrying cell wall fragments on normal microfiltration membranes. Adsorption studies and contact angle measurements have confirmed that these SUMs are net negatively charged and basically hydrophilic [153].

S-layers may also be used as matrices for covalent binding of biologically active macromolecules, such as enzymes (invertase, glucose oxidase, glucuronidase, L-glucosidase, naringinase, peroxidase), ligands (protein A, streptavidin), or mono- and polyclonal antibodies. S-layer-carrying cell wall fragments with immobilized protein A could be applied as escort particles in affinity cross-flow filtration for isolation and purification of human IgG from serum or of monoclonal antibodies from hybridoma cell culture supernatants [154, 155]. SUMs with immobilized monoclonal antibodies were also used as reaction zones for dipstick-style immunoassays [156]. For this purpose, human IgG was either directly coupled to the carbodiimide-activated carboxyl groups of the S-layer protein, or it was adsorbed onto a SUM with covalently bound protein A. Alternatively, human IgG was biotinylated and bound to a SUM onto which streptavidin was immobilized in a monomolecular layer [157]. See Figure 59.

The experimental use of crystalline bacterial surface layer proteins (S-layers) as combined carrier/adjuvants for vaccination and immunotherapy has since 1987 progressed in three areas of application: immunotherapy of cancers, antibacterial vaccines, and antiallergic immunotherapy. See Figure 60.

3.1.2. Block Copolymer Membranes

Self-assembly of reactive amphiphilic block copolymers may also be used to prepare nanostructured hydrogel membranes with exceptional permeability properties [158]. Although the block copolymer membranes (PMOXA–PDMS–PMOXA triblock) are two- to threefold thicker than conventional

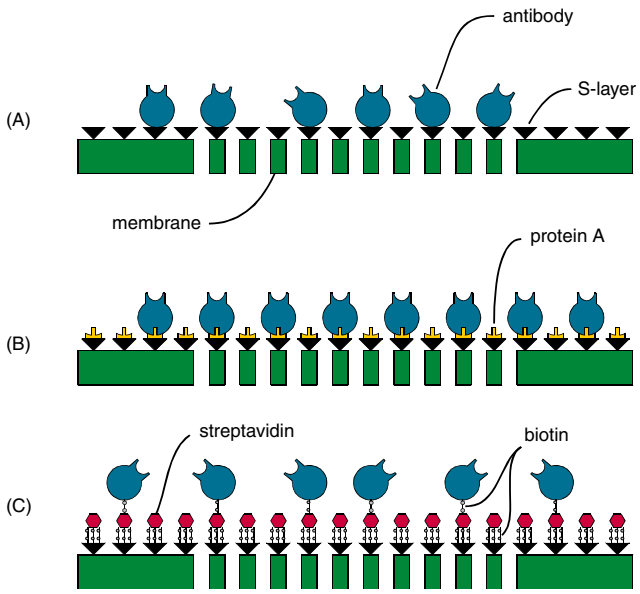


Figure 59. Schematic drawing illustrating the immobilization of IgG to carbodiimide-activated carboxyl groups of the S-layer protein of SUMs (A), to protein A covalently bound to the S-layer lattice (B), and after biotinylation to a streptavidin-modified SUM (C).

lipid bilayers, they can be regarded as mimetic of biological membranes and can further be used as a matrix or scaffold for membrane-spanning proteins [159]. Surprisingly it has been found that the proteins remain functional, despite the thickness of the membranes and even after polymerization of the reactive block copolymers [160]. The unique combination of block copolymers with membrane protein based channels allows the preparation of mechanically stable, defect-free membranes and nanovesicles [161] that have highly selective permeability and/or specific recognition sites. See Figure 61.

The block copolymer/protein hybrid shells of the nanovesicles can be regarded as a semipermeable membranes separating their internal volume from the external solution. This property opens a convenient approach to trigger the gating transition of OmpF. Large polyelectrolyte ions, such as polystyrene sulfonate, will not permeate, and

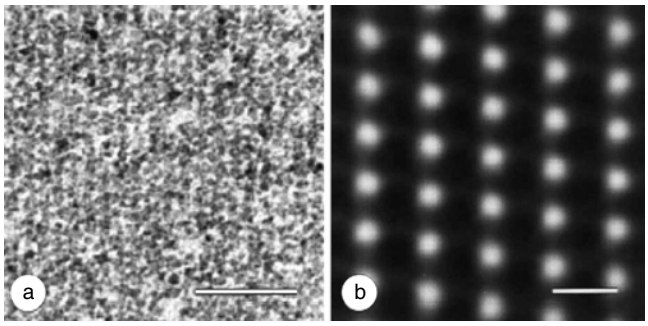


Figure 60. (a) Transmission electron micrographs of a nanometric point pattern of CdS particles obtained by biomineralization on an S-layer with oblique lattice symmetry. Protein appears white, CdS particles dark. Bar, 60 nm. (b) Corresponding computer image reconstruction to (A). Bar = 10 nm. Courtesy of *FEMS Microbiology Reviews.*

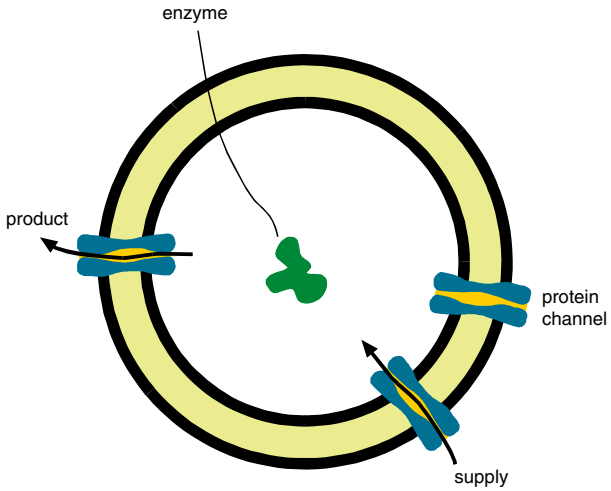


Figure 61. Schematic view of a bioreactor/nanovescicle formed by a poly-2-methyloxazoline–polydimethylsiloxane–poly-2-methyloxazoline (PMOXA-PDMS-PMOXA) triblock copolymer with encapsulated enzyme. Channel proteins in the shells of the bioreactors control the exchange of substrates and products. Reprinted with permission from [158], C. Nardin and W. Meyer, *Rev. Mol. Biotechn.* 90, 17 (2002). © 2002, Elsevier Science.

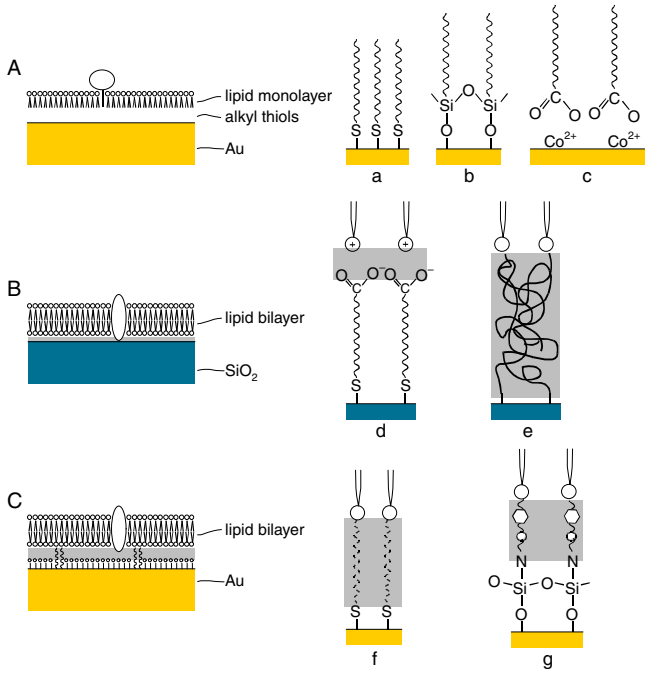


Figure 62. Architectures of supported membranes [126]. Three basic strategies are sketched here, together with some possibilities for their realization. (A) A phospholipid monolayer on a hydrophilic base layer, which in most cases consists of rigidly packed alkyl chains. This base layer may be produced by the self-assembly of alkylthiols on gold (a), of alkylsilanes on oxide surfaces (b), by the Langmuir–Blodgett (LB) transfer of a monolayer of fatty acid salts to a hydrophilic surface (c), or by the coupling of a rigid hydrophilic polymer to the surface (not shown). The phospholipid monolayer on top is either formed by entropy-driven self-assembly in solution or transferred horizontally from the air–water interface. This monolayer configuration is suited for the observation of processes on the membrane surface or for anchoring peripheral membrane proteins. (B) A phospholipid bilayer on a hydrophilic surface. This surface may be a clean glass, quartz, or silicon oxide surface on which bilayers may be formed by vesicle spreading or by two-step LB transfer. Gold surfaces must be functionalized with hydroxy-, carboxy-, or amino-terminated thiols to confer hydrophilicity; for bilayer spreading, surface charge can be exploited by using charged lipids (d). The formation of bilayers supported on a hydrophilic polymer is also possible (e). The incorporation of transmembrane proteins in such a "coating" bilayer has been demonstrated. (C) An anchored phospholipid bilayer on a hydrophilic surface. To mimic the natural situation more closely, phospholipid derivatives must be used as hydrophilic anchors for the membrane, and the interstices on the surface must either expose a hydrophilic material or be correspondingly functionalized. Anchor lipids may be thiolipids on gold (f), lipids coupled by a cross-linker to oxide surfaces (g), His-lipids, succinimidyl-lipids, etc. In this configuration, the membrane must be formed by self-assembly from vesicle suspensions or detergent solutions; it is well suited for the accommodation of transmembrane proteins. Courtesy of S. Heyse, *Biochimica et Biophysica Acta*.

therefore the sodium counterions will be distributed inside and outside the nanovesicles according to Donnan equilibrium conditions, giving rise to a Donnan potential. If this potential exceeds the critical value necessary for closure of OmpF, the substrates can no longer enter the interior of the nanoreactor vesicle (i.e., the reactors are then deactivated). The closure is a reversible process, and decreasing the potential below 100 mV may reactivate the nanoreactor vesicles. This could be carried out by diluting the system with buffer or by increasing the Na^+ concentration in the system.

Peptide nanochannels or nanotubes have also been formed by cyclo([L-Trp-D-Leu]$_3$-L-Gln-D-Leu) in phospholipid multilayers obtained by air drying on a substrate [162]. Nanotubes formed from cyclo([L-Trp-D-Leu]$_4$) have also been allowed to assemble into preformed monolayers of dodecanethiol or octadecyl sulfide deposited on gold films [163].

Bonding of the bilayer can be adjusted by coating of the membrane using a matching precoat, for instance high-quality alkyl monolayers. The stability of the precoat is of very high importance for the final stability of the total stack [164, 165]. See Figure 62.

3.2. Biosensors

Micro- and nanoengineering technologies offer revolutionary possibilities for biosensors and sensor arrays for drug and chemical screening and environmental monitoring. When a cell membrane detects (senses) a target molecule, it can turn electrical currents on or off by opening or closing molecular channels. When the channels are open, charged ions can pass in or out of the cell. These ions would not pass through the otherwise insulating membrane. When these channels open, the ion flow creates a potential difference across the membrane, which in turn creates a current. Such biosensors have a huge range of potential uses, especially in medicine, for detecting drugs, hormones, viruses, and pesticides and to identify gene sequences for diagnosing genetic disorders. The ability to detect the modulation of cell membrane channel activity by the binding of therapeutic agents is considered crucial for rational and efficient drug discovery and design. Since combinatorial libraries of potential therapeutic compounds are rapidly growing, fast and highly

sensitive methods for functional drug screening are required. An attractive possibility is the use of self-assembled tethered membranes containing specific channel receptors as the sensing element in an otherwise solid-state biosensing device. Massive arrays of individually addressable microsensors with integrated fluid handling are conceivable. Even very simple sensor designs offer valuable advances in low-cost sensing for clinical medicine and the food and hygiene sectors. The much-discussed "artificial nose" containing a dense array of receptor sites affording unambiguous identification of molecular species could analyze the breath of patients for known chemical signatures of diseases such as liver cirrhosis and lung cancer.

3.2.1. Scaffolds

Microengineered membranes coated with titanium with a pore size of 500 nm have also been used [166] as a scaffold for a functional monolipid layer obtained by the coalescence and spreading of corresponding lipid vesicles on the membrane surface.

Micro- and nanoengineered membranes can further be used as a scaffold for the construction of functional 2D nanoarchitectures to probe, for example, selective molecular transport and controlled release properties of biomedical molecules. See Figure 63.

Cornell et al. [167] have elaborated tethered supported lipid bilayers into devices that will sense both large and small analytes. The impedance of the bilayers depends upon the density of open gramicidin channels. In one manifestation, an analyte (such as a protein) with two antigenic sites binds to antibodies at the membrane surface disrupting a preexisting channel with the disadvantage that the response is a diminution of a large preexisting signal. In a second manifestation, channels are opened when an analyte, which can in this case be a small molecule, releases half channels from an immune complex in the upper leaflet of the bilayer so that they can find half channels in the lower leaflet. Another promising approach is stochastic sensing with engineered pores [168]. Stochastic sensing is based on the detection of individual binding events between analyte molecules and a single pore. The read-out is the single-channel electrical current. The frequency of the binding events is determined by the concentration of the analyte. The nature of the binding events (e.g., the magnitude and duration of the associated signal) is determined by the properties of the analyte. The ability to identify an analyte by its characteristic signature is a distinctive feature of stochastic sensing. Engineered versions of α-hemolysin were first used to detect and quantify divalent metal ions.

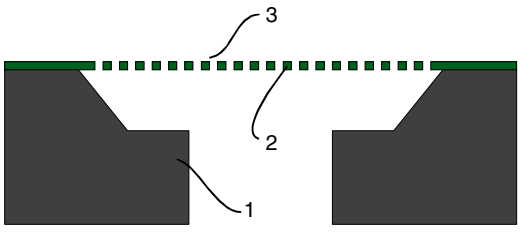


Figure 63. Support structure 1 with a nanoengineered scaffold 2 for a functional membrane layer 3.

4. NANOTUBES

4.1. Introduction

Nanotubes can be defined as hollow cylindrical structures and constructs with an inner diameter smaller than 100 nm. The current worldwide research is mainly focused on:

- *Peptide nanochannels and tubes* that can be found in cell membranes acting as size or charge selective barriers between the inner an outer lipid bilayer cell membrane and can also be artificially constructed and be used as molecule specific biosensors (see the section? on nanoengineered biomembranes).
- *Carbon and carbonlike dense and open nanotubes* that can be grown selectively on appropriate prepared substrates and can be used for molecule specific adsorption/separation processes or for electrical nanowiring applications.
- *Templated nanotubes or tubules* in which, for example, an anodized alumina membrane is used as a template for the deposition of appropriate materials (carbon, gold, etc.) inside the long cylindrical pores (diameter 50–200 nm) of the alumina membrane. Procedure: immerse an alumina membrane in a solution of organometallic nickel, evaporate solvent, leaving Ni film in pores, form nanotubes in membrane pores via CVD of carbon from ethylene decomposition [169]. Increasing deposition time yields solid carbon nanofibersa rather than hollow tubes. Recently silica nanotubes [170, 171] have also been obtained. Other tubular structures formed by heavier element compounds have been predicted, such as GaSe [172], and synthesized, such as WS_2 and MoS_2 [173].

4.2. Carbon Nanotubes

Two types of carbon nanotubes exist: those originally observed by Iijima [174, 175] were multiwall nanotubes, formed by concentric shells of apparently seamless cylinders of graphene, having a separation between them similar to that in graphite. More recently, single-wall nanotubes (SWNTs) have also been synthesized. As their name indicates, these consist of a single seamless cylinder of graphene [176]. Soon after the discovery of carbon nanotubes it was proposed that other compounds forming graphite-like structures, such as BN [177], BC_3 [178], BC_2N, and CN [179], could also form nanotubular structures and indeed BN, BC_3, and BC_2N have already been synthesized [180]. SWNTs were first synthesized in an arc discharge in the presence of a transition metal catalyst [181, 182]. Since then significant efforts have been directed at optimizing conditions for the arc production of SWNTs [183, 184]. For example, a modification of the arc discharge process allowing production of high-quality nanotube material within a restricted region of the apparatus has been reported [185]. However, the overall nanotube yield for this process remains relatively low. Other methods for SWNT synthesis have also been developed. An efficient SWNT production method is the laser vaporization of graphite/transition metal catalyst targets in a heated oven, which can produce nanotube yields of $>70\%$. SWNT syntheses by metal-catalyzed disproportionation of carbon monoxide and catalytic decomposition of acetylene have also been reported [186].

Ren et al. [187] discovered that vertically aligned multiwall carbon nanotubes could also be grown using plasma enhanced chemical vapor deposition (PECVD). Other plasma techniques nowadays are hot filament PECVD [187], dc (glow discharge) PECVD [188], microwave PECVD [189], and inductively coupled plasma PECVD [190]. See Figure 64.

Structural studies with transmission electron/scanning electron microscopes revealed that these nanofibersa contained the catalyst at their tips which suggests a tip-growth mechanism. Over a large range of diameters (30–400 nm), the structures were hollow (i.e., large inner diameters, hence "tube"like) with graphene planes running parallel to the length axis, and bamboo-type defects along its length.

The field of prospective applications of carbon nanotubes continues to expand in nanofluidics and nanofiltration applications. Carbon nanotubes might be used as a nanopipets [192], as sieves for DNA sequencing applications, for enantiomer separation, etc. A fundamental issue is the ability of a fluid to wet the interior of a CNT as this would facilitate solution chemistry inside CNTs [193]. The surface of pure graphite is hydrophilic; the contact angle of water on graphite is 80–90° [194]. Different studies reveal different results [195–197] on the wettability of pure carbon nanotubes. The addition of impurities during the manufacturing of the nanotubes or surface modification steps afterward might be a solution to overcome serious wetting problems. See Figure 65.

4.3. Templated Nanotubes and Nanofiltration

Classical filtration membranes and in particular anodized alumina membranes [199] and nanosieves [200] may also be transformed to nanotubular structures by deposition of appropriate materials in the filtration pores or channels. They will enable new ways of separating and detecting analytes for applications in chiral separations and single-molecule sensing. See Figure 66.

Nanotubule membranes are easy to make and each pore in the membrane essentially is a nanobeaker in which chemistry can occur. Metals can be deposited inside the pores, either electrochemically or through so-called electroless plating by using a reducing agent to plate the metal from the solution. Nanotubules of inorganic materials such as silica or titania can be prepared through sol–gel chemistry, and carbon nanotubules can be made through chemical vapor deposition of ethylene inside the pores. See Figure 67.

The nanotubules come out embedded in a membrane, aligned, and monodisperse—that is, of uniform diameter and of a uniform length equal to the thickness of the membrane. Depending on deposition reaction times, thin- or thick-wall tubules are formed. The tubules can be capped on both ends to create a cluster of individual confined

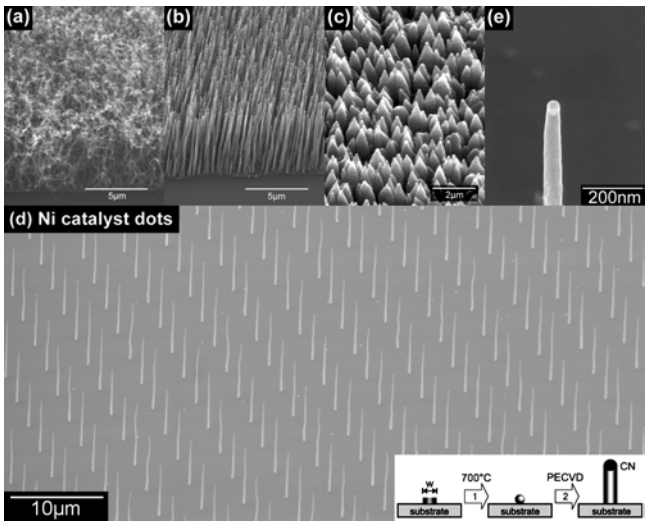


Figure 64. The possibilities of the dc PECVD process. A substrate containing thin film Ni catalyst is exposed to C_2H_2 and NH_3 gases at 700 °C [i.e., standard thermal CVD, (a)] with the exception that the substrate is biased at −600 V (i.e., to initate dc glow discharge). The degree of nanotube/nanofiber vertical alignment depends on the magnitude of the voltage applied to the substrate during growth and −600 V produced the straight structures of (b). The thickness of the thin film Ni catalyst was found to determine the average diameter and inversely the length of the nanofibersa. The yield and density of the nanofibersa were controlled by the use of different diffusion barrier materials under the Ni catalyst. The shape of the structures could be varied from very straight nanotubelike to conical tiplike nanofibersa by increasing the ratio of C_2H_2 in the gas flow (c). Example of patterned growth of CN (d). Close-up of vertically grown nanotube (e). Reprinted with permission from [191], K. B. K. Teo et al., University of Cambridge, UK.

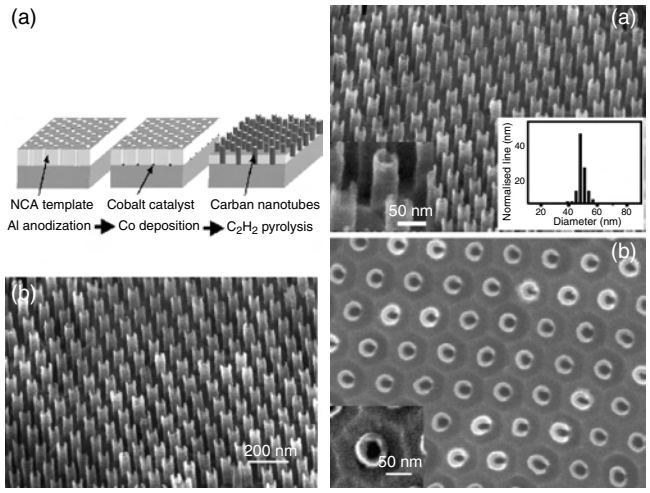


Figure 65. Highly ordered carbon nanotube arrays. Top left: Process scheme of fabrication. Bottom left: SEM image of the resulting hexagonally ordered array of carbon nanotubes fabricated using the method in (a). Top right: SEM image showing oblique view of periodic carbon nanotube array. The inset at the lower left is an enlarged view of the tubes. The inset at the lower right is a histogram of the nanotube diameter showing a narrow size distribution around 47 nm. Top view SEM image, bottom right: carbon nanotubes showing hexagonal close-packed geometry. The hexagonal cells have sides approximately 57 nm long and the intercell spacing is 98 nm. The slight splitting of the tube ends and the apparent increase in tube wall thickness is an artifact of the nonspecialized ion-milling apparatus that was used in the experiments. The inset shows a close-up view of a typical open-ended carbon nanotube in its hexagonal cell. Reprinted with permission from [198], J. Li et al., *Appl. Phys. Lett.* 75, 367 (1999). © 1999, American Institute of Physics.

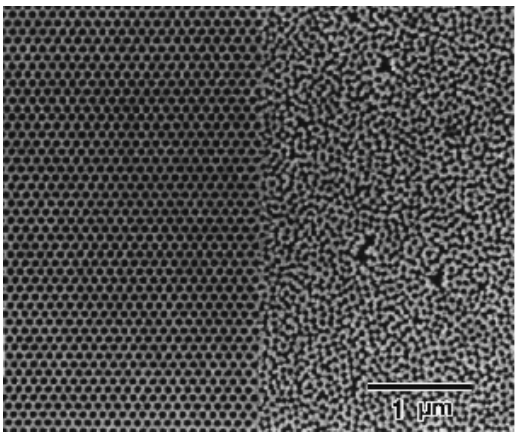


Figure 66. SEM micrograph of ordered nanochannels in anodic porous alumina. The anodizing of Al was conducted under a constant voltage condition in an oxalic acid solution. The hole interval of anodic porous alumina, in other words, the cell size, was determined by the applied voltage used for anodization. It was reported that the cell size has a good linear relationship with the applied voltage, where the proportionality constant of cell size for a specific applied voltage is approximately 2.5 nm/V. In the case of the sample shown, anodizing was conducted under a constant voltage of 40 V after pretexturing of a 100 nm period. Ideally ordered hole development was observed only in a pretextured area (left), while random development of holes took place in the untreated area (right). The hole interval corresponded to that of the pretextured concaves of Al. From this result, it was concluded that concave features formed by indentation could act as an initiation point and guide the growth of channels. Reprinted with permission form [201], H. Asoh et al., *J. Electrochem. Soc.* 148, 152 (2001). © 2001, The Electrochemical Society.

spaces, or the membrane can be removed to release individual nanotubules. The generality of this template-based synthetic methodology broadens the scope of nanomaterials that can be prepared. One of the ealiest separation applications to be explored was selective-ion transport. A simple experiment with a gold nanotubule membrane in a U-tube cell demonstrates this phenomenon [202]. On one side of the membrane is a feed solution containing KCl, which is colorless, and the cationic dye methylene blue. On the other side is a receiver solution of KCl. After a while, the colorless receiver solution turns blue, indicating transport of the cationic dye across the membrane. But when the feed contains permanganate anion, which is red, the receiver

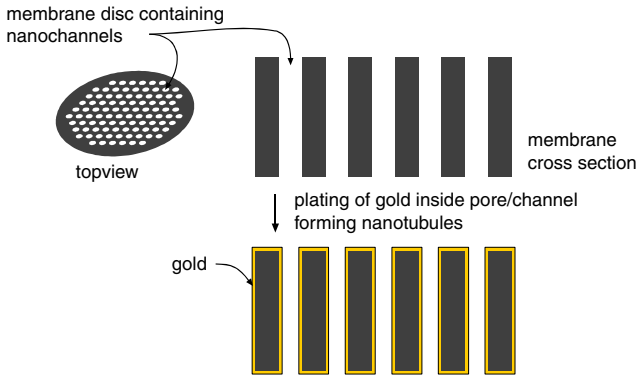


Figure 67. Schematic overview of the formation of nanotubules.

remains colorless. The anion, although smaller than the cationic dye, cannot cross the membrane. Only cations pass because the nanotubule walls have excess negative charge due to adsorbed chloride ions. The excess charge can also be controlled by applying a potential to the membrane provided that the membrane material is electrically conductive. If a positive potential is applied, the nanotubule wall will have excess positive charge and the nanotubule membrane will reject cations and transport only anions. At negative applied potentials, the membrane will transport cations and reject anions. At the specific potential where the wall is electrically neutral, the membrane will be nonselective. Thus, the nanotubules make up a switchable ion-exchange membrane [203]. Having shown these membranes to have charge- and size-based selectivities, attempts were made to engender chemical selectivity by modifying the chemistry of the nanotubule walls in which alkyl and other groups are attached to the nanotubule walls through gold–sulfur bonds [204, 205]. Cysteine is a thiol as well as an amino acid. So it has both amino and carboxyl groups. At low pH, it will be positively charged; at high pH, it will be negatively charged. Lee and Martin have shown [206] that the cysteine-decorated membrane rejects cations at low pH and anions at high pH.

Chiral separation is important in the pharmaceutical industry, as it has become clear that the enantiomers of drugs produced as racemic mixtures could have very different properties. Lakshmi et al. [207] used a "sandwich" assembly to separate D-phenylalanine from L-phenylalanine. Kobayashi et al. [208] achieved the highly sensitive detection at 10^{-11} M with unadorned gold nanotubule membranes through which an ionic current passes.

GLOSSARY

Biomembrane Natural membrane consisting of a lipid bilayer and certain protein channels for specific ion exchange.

Biosensor Device to measure chemical composition of liquid or gas using a cell membrane. Especially for use in drug screening and health and environmental monitoring.

Chemical vapor deposition (CVD) Chemical vapor deposition, a process to form thin layers of specific solid material from a gas phase.

Filtration (ultra, nano, micro) Selective permeation of any kind of molecule, bacteria, micro-organism or particle through a device, usually a thin membrane using the properties of the membrane (most basic on size selection, but more specific by difference in affinity with different molecules).

Laser interference lithography (LIL) Light exposure method using two coherent laser beams, thereby obtaining a regular line pattern. With LIL line patterns with a typical width down to 100 nm can easily be achieved. With multiple exposures, dot or hole patterns can be formed.

Lipid bilayer Natural dense membrane made from two long molecules, each with a polar head and a nonpolar tail. The molecules are arranged such that the nonpolar tails are in line with each other. By this the membrane has a hydrophilic inside and thus acts as a barrier for ions and hydrophilic liquids.

Membrane Boundary used to separate two phases. Depending on the properties of the membrane, selective transport can take place between the two phases through the membrane.

Micromachining Technique originating from the semiconductor technology used to fabricate mechanical structures and devices with typical sizes between 0.1 and 100 micrometers. The technique started with silicon material, but nowadays also polymers are being used.

Microsieve Thin, flat membrane made of silicon nitride with silicon support, having pores in a regular pattern with a very uniform pore size. Using micromachining, the pore size can be adjusted according to the application. Due to the thin membrane layer and high porosity, the flow resistance is very small.

Nanocontact printing Patterning method used to print inks, proteins, or any kind of liquid with the use of (porous) stamps containing a pattern with nanodimensions.

Nanolithography Exposure method to form patterns in light sensitive lacquers having a resolution well below 500 nm. Techniques that can be used are e-beam, ion beam, X-ray (synchrotron), or laser interference lithography.

Nanotubes Fiber structures with a diameter less than 200 nm. Due to the construction of the tube (a highly ordered molecular structure), the tubes are used for research toward new materials (strong fibers, highly conductive materials).

Phase inversion Process to form a solid material out of a liquid mixture of a polymer and a solvent/nonsolvent.

Photonic crystal Ordered regular structure with a lattice dimension in the range of the wavelength of light. By locally altering the lattice, light can be propagated or blocked.

Self-assembled monolayer (SAM) A single molecule thick layer that is arranged highly ordered without using special patterning techniques. The structure of the molecule itself organizes the position of the molecule. Used for thin layers with a special affinity to certain materials.

REFERENCES

1. Osmosis was discovered by Abbé J. A. Nollet in 1748 using a pig's bladder membrane.
2. S. Loeb and S. Sourirajan, *Adv. Chem. Ser.* 38, 117 (1981).
3. C. J. M. van Rijn, Membrane Filter as well as a Method of Manufacturing the Same, PCT Application 95/1386026.
4. S. Chien, *Microvascular Res.* 44, 243 (1993).
5. M. H. V. Mulder, "Basic Principles of Membrane Technology." Kluwer Academic, Dordrecht, 1991.
6. I. Eriksson, *Env. Progress* 7, 58 (1988).
7. W. J. Conlon, *Desalination* 56, 203 (1985).
8. J. Cadotte, R. Forester, M. Kim, R. Petersen, and T. Stocker, *Desalination* 70, 77 (1988).
9. X.-L. Wang, T. Tsuru, M. Togoh, S. I. Nakao, and S. Kimura, *J. Chem. Eng. Japan* 28, 186 (1995).
10. X.-L. Wang, T. Tsuru, M. Togoh, S. I. Nakao, and S. Kimura, *J. Chem. Eng. Japan* 28, 372 (1995).
11. W.-R. Bowen and H. Mukthar, *J. Membrane Sci.* 112, 263 (1996).
12. M. Kurihara and Y. Himeshima, *Polym. J.* 23, 513 (1991).
13. L. P. Raman, M. Cheryan, and N. Rajagopalan, *Chem. Eng. Progr.* 3, 68 (1994).
14. M. Dekker and R. Boom, *TIBTech* 13, 129 (1995).
15. R. J. Petersen, *J. Membrane Sci.* 83, 81 (1993).
16. J. K. Mitchell, *Am. J. Med. Sci.* 25, 100 (1833).
17. T. Graham, *Philos. Mag.* 32, 401 (1866).
18. J. Haggin, *Chem. Eng.* 7 (1988).
19. A. Bos, R. T. Chern, and C. N. Provan, *Macromolecules* 24, 2203 (1991).
20. J. H. Petropoulos, *J. Membrane Sci.* 75, 47 (1992).
21. M. Wessling, S. Schoeman, Th. v.d. Boomgaard, and C. A. Smolders, *Gas Sep. Purif.* 5, 222 (1991).
22. W. J. Koros and I. Pinnau, "Polymeric Gas Separation Membranes," p. 209. CRC Press, Baca Raton, FL, 1994.
23. J. K. Mitchell, *Am. J. Med.* 13, 36 (1830).
24. K. N. Mani, *J. Membr. Sci.* 58, 117 (1991).
25. G. Pourcelly, "Handbook on Bipolar Membrane Technology." Twente Univ. Press, Twente, 2000.
26. S. Koter and A. Warszawski, *Polish J. Env. Stud.* 9, 45 (2000).
27. Y. Mizutani, *J. Membr. Sci.* 49, 121 (1990).
28. T. A. Davis, J. D. Genders, and D. Pletcher, "A first course in Ion Permeable Membranes." The Electrochemical Consultancy, Romsey, UK, England, 1997.
29. B. E. Yoldas, *Ceram. Bull.* 54, 289 (1975).
30. B. Sea and K.-H. Lee, *Bull. Kor. Chem. Soc.* 22, 1400 (2001)
31. Y.-S. Lin, C. H. Chang, and R. Gopalan, *Ind. Chem. Eng. Res.* 33, 860 (1994).
32. C. Guizard et al., *Mater. Sci. Forum* 154, 152 (1994).
33. R. S. A. de Lange, J. H. A. Hekkink, K. Keizer, and A. Burggraaf, *J. Membrane Sci.* 95, 57 (1995).
34. N. K. Raman and C. Brinker, *J. Membrane Sci.* 105, 273 (1995).
35. R. M. Boom, T. van den Boomgaard, and C. A. Smolders, *J. Membrane Sci.* 90, 231 (1994).
36. R. M. Boom, T. van den Boomgaard, and C. A. Smolders, *Macromolecules* 27, 2034 (1994).
37. Z.-K. Xu, M. Böhning, J. Springer, N. Steinhauser, and R. Mülhaupt, *Polymer* 38, 581 (1997).
38. Z.-K. Xu, M. Böhning, J. D. Schultze, G.-T. Li, J. Springer, F. P. Glatz, and R. Mülhaupt, *Polymer* 38, 1573 (1997).
39. T.-S. Chung, E. R. Kafchinski, and P. Foley, *J. Membr. Sci.* 75, 181 (1995).
40. S. A. Stern, Y. Liu, and W. A. Fed, *J. Polym. Sci.* 31, 939 (1993).
41. M. Smaihi, J.-C. Schrotter, C. Lesimple, I. Prevost, and C. Guizard, *J. Membr. Sci.* 161, 157 (1999).
42. L. M. Robeson, *J. Membr. Sci.* 62, 165 (1991).
43. G. Maier, *Angew. Chem. Int. Ed.* 37, 2960 (1998).
44. Z.-K. Xu, Y.-Y. Xu, and M. Wang, *J. Appl. Polym. Sci.* 69, 1403 (1998).
45. K. Kusakabe, K. Ichiki, J. Hayashi, H. Maeda, and S. Morooka, *J. Membr. Sci.* 115, 65 (1996).
46. T. C. Merkel, B. D. Freeman, R. J. Spontak, Z. He, I. Pinnau, P. Meakin, and A. J. Hill, *Science* 296, 519 (2002).
47. H. V. Jansen, Plasma Etching in Micro-technology, Thesis, 1996.
48. C. J. M. van Rijn and M. C. Elwenspoek, *IEEE Proc. MEMS* 83 (1995).
49. M. Rossi, Mikrostrukturierung, in "Vortrag zu den Übungen in Halbleiterphysik II," Stuttgart, June 1999.
50. Pedrotti, Pedrotti, Bausch, and Schmidt, "Optik, Eine Einführung." Prentice–Hall, Munich, 1996.
51. D. Meschede, "Optik, Licht und Laser." Teubner, Stuttgart/Leipzig, 1999.
52. D. Widmann, H. Mader, and H. Friedrich, "Technologie hochintegrierter Schaltungen (2. Auflage)." Springer-Verlag, Berlin/Heidelberg/New York, 1996.
53. A. Kroyan, J. Bendik, O. Semprez, N. Farrar, C. Rowan, and C. A. Mack, *Proc. SPIE* 4000 (2000).
54. C. A. Mack, "Inside PROLITH: A Comprehensive Guide to Optical Lithography Simulation." FINLE Technologies, 1997.
55. E. Hecht, "Optik (2. Auflage)." Oldenbourg-Verlag, München/Wien, 1999.

56. M. Young, "Optik, Laser, Wellenleiter." Springer-Verlag, Berlin/Heidelberg/New York, 1997.
57. P. W. Hawkes and E. Kasper, "Principles of Electron Optics." Academic Press, London, 1989.
58. P. Grivet, "Electron Optics." Elsevier, Oxford, 1965.
59. M. Hatzakis, *J. Electrochem. Soc.* 116, 1033 (1969).
60. C. P. Umbach, C. Van Haesendonck, R. B. Laibowitz, S. Washburn, and R. A. Webb, *Phys. Rev. Lett.* 56, 386 (1986).
61. V. Chandrasekhar, M. J. Rooks, S. Wind, and D. E. Prober, *Phys. Rev. Lett.* 55, 1610 (1985).
62. B. J. van Wees, H. van Houten, C. W. J. Beenakker, J. G. Williamson, L. P. Kouwenhoven, D. van der Marel, and C. T. Foxon, *Phys. Rev. Lett.* 60, 848 (1988).
63. M. J. Rooks, C. C. Eugster, J. A. del Alamo, G. Snider, and E. Hu, *J. Vac. Sci. Technol. B* 9, 2856 (1991).
64. P. H. Woerlee, G. A. M. Hurkx, W. J. M. J. Josquin, and J. F. C. M. Verhoeven, *Appl. Phys.* Lett. 47, 700 (1985).
65. E. Anderson, V. Boegli, M. Schattenburg, D. Kern, and H. Smith, *J. Vac. Sci. Technol. B* 9, 3606 (1991).
66. R. Viswanathan, D. Seeger, A. Bright, T. Bucelot, A. Pomerene, K. Petrillo, P. Blauner, P. Agnello, J. Warlaumont, J. Conway, and D. Patel, *J. Vac. Sci. Technol. B* 11, 2910 (1993).
67. S. Y. Chou, H. I. Smith, and D. A. Antoniadis, *J. Vac. Sci. Technol. B* 4, 253 (1986).
68. S. Austin and F. T. Stone, *J. Appl. Opt.* 15, 1071 (1976).
69. B. de A. Mello, I. F. da Costa, C. R. A. Lima, and L. Cescato, *Appl. Opt.* 34, 597 (1995).
70. L. Mashev and S. Tonchev, *Appl. Phys. A* 26, 143 (1981).
71. C. J. M. van Rijn, W. Nijdam, S. Kuiper, G. J. Veldhuis, H. van Wolferen, and M. Elwenspoek, *J. Micromech. Microeng.* 9, 170 (1999).
72. E. H. Anderson, C. M. Horwitz, and H. I. Smith, *Appl. Phys. Lett.* 43, 874 (1983).
73. H. Z. Saleem and S. R. J. Brueck, *J. Vac. Sci. Technol. B* 11, 658 (1993).
74. C. J. M. van Rijn, G. J. Veldhuis, and S. Kuiper, *Nanotechnology* 9, 343 (1998).
75. C. J. M. van Rijn, W. Nijdam, S. Kuiper, G. J. Veldhuis, H. van Wolferen, and M. Elwenspoek, *J. Micromech. Microeng.* 9, 170 (1999).
76. A. Fernandez, P. J. Bedrossian, S. L. Baker, S. P. Vernon, and D. R. Kania, *IEEE Trans. Mag.* 32, 4472 (1996).
77. C. O. Boltzer, C. T. Harris, S. Rabe, D. D. Rathman, M. A. Hollis, and H. I. Smith, *J. Vac. Sci. Technol. B* 12, 629 (1994).
78. J. G. E. Gardeniers, H. A. C. Tilmans, and C. G. C. Visser, *J. Vac. Sci. Technol. A* 14, 2879 (1996).
79. Y. Xia and G. M. Whitesides, *Angew. Chem. Int. Ed.* 37, 550 (1998).
80. M. C. Porter,"Handbook of Industrial Membrane Technology." Noyes, Park Ridge, NJ, 1990.
81. "Handbook of Chemistry and Physics," 1st stud. ed. CRC Press, Baca Raton, FL, 1988.
82. C. J. M. van Rijn, M. van der Wekken, W. Nijdam, and M. C. Elwenspoek, *J. Microelectromech. Syst.* 6, 48 (1997).
83. H. V. Jansen, Plasma Etching in Micro-technology, Thesis, University of Twente, The Netherlands, 1996.
84. Oxford Instruments, Plasma Technology, North End, Yatton, Bristol BS19 4AP, England.
85. J. A. Theil, *J. Vac. Sci. Technol. B* 13, 2145 (1995).
86. H. V. Jansen, M. J. de Boer, H. Wensink, B. Kloeck, and M. C. Elwenspoek, *Microsyst. Technol.* 6 (2000).
87. A. Hölke and H. T. Henderson, *J. Micromech. Microeng.* 9, 51 (1999).
88. P. Krause and E. Obermeier, *J. Micromech. Microeng.* 5, 112 (1995).
89. H. Seidel, L. Csepregi, A. Heuberger, and H. Baumgärtel, *J. Electrochem. Soc.* 137, 3612 (1990).
90. G. J. Burger, E. J. T. Smulders, J. W. Berenschot, T. S. J. Lammerink, J. H. J. Fluitman, and S. Imai, *Proc. Transducers* 573 (1995).
91. B. de Heij, MSc. Thesis, University of Twente, 1997.
92. J. Brugger, J. W. Berenschot, S. Kuiper, W. Nijdam, B. Otter, and M. Elwenspoek, *Microelectron. Engi.* 53, 403 (2000).
93. M. Kölbel, R. W. Tjerkstra, J. Brugger, C. J. M. van Rijn, W. Nijdam, J. Huskens, and D. N. Reinhoudt, *Nanolett.* 2, 1339 (2002).
94. E. Yablonovitch and T. J. Gmitter, *Phys. Rev. Lett.* 63, 1950 (1989).
95. K. Inoue, M. Wada, K. Sakoda, A. Yamanaka, M. Hayashi, and J. W. Haus, *Jpn. J. Appl. Phys.* 33, 1463 (1994).
96. A. Rosenberg, R. J. Tonucci, H.-B. Lin, and A. J. Campillo, *Opt. Lett.* 21, 830 (1996).
97. J. D. Joannopoulos, R. D. Meade, and J. N. Winn, "Photonic Crystals, Molding the Flow of Light." Princeton Univ. Press, Princeton, NJ, 1995.
98. E. Yablonovitch, *Phys. Rev. B* 58, 2059 (1987).
99. L. Vogelaar, W. Nijdam, H. A. G. M. van Wolferen, R. M. de Ridder, F. B. Segerink, E. Flück, L. Kuipers, and N. F. van Hulst, *Adv. Mater.* 13, 1551 (2001).
100. M. Lončar, T. Doll, J. Vučković, and A. Scherer, *J. Lightw. Technol.* 18, 1402 (2000).
101. T. F. Krauss, R. M. De La Rue, and S. Brand, *Nature* 383, 699 (1996).
102. S. G. Johnson, S. Fan, P. R. Villeneuve, and J. D. Joannopoulos, *Phys. Rev. B.* 60, 5751 (2000).
103. See for example, M. D. B. Charlton, M. E. Zoorob, G. J. Parker, M. C. Netti, J. J. Baumberg, S. Cox, and H. Kemhadjian, *Mater. Sci. Eng. B* 74, 17 (2000).
104. V. Berger, O. Gauthier-Lafaye, and E. Costard, *Electron. Lett.* 33, 425 (1997).
105. C. Peeters, E. Flück, A. M. Otter, M. L. M. Balistreri, J. P. Korterik, L. Kuipers, and N. F. van Hulst, *Appl. Phys. Lett.* 77, 142 (2000).
106. H. Masuda, M. Ohya, K. Nisahio, H. Asoh, M. Nakao, M. Nohtomi, A. Yokoo, and T. Tamamura, *Jpn. J. Appl. Phys.* 39, 1039 (2000).
107. H. Masuda, M. Ohya, H. Asoh, M. Nakao, M. Nohtomi, and T. Tamamura, *Jpn. J. Appl. Phys.* 38, 1403 (1999).
108. H. Verweij, A. Nijmeijer, and A. ten Elshof, STW Project, Inorganic Material Science, University Twente.
109. X. Yang, J. M. Yang, Y. Tai, and C. Mo, *Sensors Actuators* 73, 184 (1999).
110. A. J.Frank, K. F. Jensen, and M. A. Schmidt, *Proc. IEEE Conf. MEMS* 382 (1999).
111. T. A. Desai, D. J. Hansford, and M. Ferrari, *Biomolecular Engi.* 17, 23 (2000).
112. T. A. Desai, *Med. Eng. Phys.* 1 (2000).
113. T. A. Desai, D. J. Hansford, L. Leoni, M. Essenpreis, and M. Ferrari, *Biosensors Bioelectron.* 15, 453 (2000).
114. M. Essenpreis, T. A. Desai, M. Ferrari, and J. Hansford, Implantable Analyte Sensor, PCT Patent Application.
115. C. J. M. van Rijn, W. Nijdam, L. A. V. G. van der Stappen, O. J. A. Raspe, L. Broens, and S. van Hoof, *Proc. EBC Congress* 501 (1997).
116. D. S. Ryder, C. R. Davis, D. Anderson, F. M. Glancy, and G. N. Power, *MBAA Tech. Quart.* 25, 67 (1988).
117. Q. Gan, R. W. Field, M. R. Bird, R. England, J. A. Howell, M. T. Mckechnie, and C. L. O'Shaughnessy, *Trans. Inst. Chem. Eng.* 75 (1997).
118. P. J. I. Janssen, Replacement of Kieselguhr Filtration by High Flux Microsieves, Report of research at the Grolsch Brewery, Groenlo, 1997.
119. P. Banplain-Avet, N. Doubrovine, C. Lafforgue, and M. Lalande, *J. Mem. Sci.* 152, 151 (1999).

120. B. Czech, *The Brewer* 103 (1995).
121. S. Kuiper, C. J. M. van Rijn, W. Nijdam, and O. Raspe, *J. Membrane Sci.* 196, 159 (2002).
122. S. M. Bezrukov, I. Vodyanoy, and V. A. Parsegian, *Nature* 370, 279 (1994).
123. C. Nardin, J. Widmer, M. Winterhalter, and W. Meier, *J. Phys. E* 4, 403 (2001).
124. M. Winterhalter, C. Hilty, S. M. Bezrukov, C. Nardin, W. Meier, and D. Fournier, *Talanta* 55, 965 (2001).
125. H. Bayley, *Current Opinion Biotechnol.* 10, 94 (1999).
126. S. Heyse, T. Stora, E. Schmid, J. H. Lakey, and H. Vogel, *Biochim. Biophys. Acta* 85507, 319 (1998).
127. P. van Gelder, F. Dumas, and M. Winterhalter, *Biophys. Chem.* 85, 153 (2000).
128. C.-H. Heldin and M. Purton, "Signal Transduction." Chapman and Hall, London, 1996.
129. J. T. Hancock, "Cell Signalling." Addison–Wesley Longman, Essex, 1997.
130. T. Schirmer, T. A. Keller, Y. F. Wang, and J. P. Rosenbusch, *Science* 267, 512 (1995).
131. N. Saint, K. L. Lou, C. Widmer, M. Luckey, T. Schirmer, and J. P. Rosenbusch, *J. Biol. Chem.* 271, 20676 (1996).
132. S. M. Bezrukov and M. Winterhalter, *Phys. Rev. Lett.* 85, 202 (2000).
133. B. Schuster, D. Pum, O. Braha, H. Bayley, and U. B. Sleytr, *Biochim. Biophys. Acta* 1370, 280 (1998).
134. S. M. Bezrukov, I. Vodyanoy, and V. A. Parsegian, *Nature* 370, 279 (1994).
135. S. M. Bezrukov, *J. Membr. Biol.* 174, 1 (2000).
136. U. B. Sleytr, P. Messner, D. Pum, and M. Sara, "Crystalline Bacterial Cell Surface Proteins," p. 1. Academic Press, Austin, TX, 1996.
137. P. Messner and U. B. Sleytr, *Adv. Microbial Physiolo.* 33, 213 (1992).
138. T. J. Beveridge, *Curr. Opin. Struct. Biol.* 4, 202 (1994).
139. U. B. Sleytr and P. Messner, "Electron Microscopy of Subcellular Dynamics," p. 13. CRC Press, Boca Raton, FL, 1989.
140. U. B. Sleytr, H. Bayley, and M. Sára, *FEMS Microbiol. Rev.* 20, 151 (1997).
141. D. Pum and U. B. Sleytr, "Crystalline Bacterial Cell Surface Proteins," p. 175. Academic Press, Austin, TX, 1996.
142. M. Montal and P. Muëller, *Proc. Natl. Acad. Sci. USA* 69, 3561 (1972).
143. H. Schindler, *Methods Enzymol.* 171, 225 (1989).
144. W. Hanke and W.-R. Schlue, "Biological Techniques Series," Vol. 9. Academic Press, London, 1993.
145. M. Sára and U. B. Sleytr, *J. Bacteriol.* 169, 4092 (1987).
146. M. Sára and U. B. Sleytr, *J. Membrane Sci.* 33, 27 (1987).
147. M. Sára, S. Kuëpcuë, and U. B. Sleytr, "Crystalline Bacterial Cell Surface Proteins" (R. G. Landes, Ed.), p. 133. Academic Press, Austin, TX, 1996.
148. M. Sára and U. B. Sleytr, *Micron* 27, 141 (1996).
149. S. Nakao, *J. Membrane Sci.* 96, 131 (1994).
150. K. C. Schuster, H. F. Mayer, R. Kieweg, W. A. Hampel, and M. Sára, *Biotechnol. Bioeng.* 48, 66 (1995).
151. H. Kuhn, U. Friederich, and A. Fiechter, *Appl. Microbiol. Biotechnol.* 6, 341 (1979).
152. U.B. Sleytr and M. Sára, US Patent 4, 886, 604, 1989.
153. S. Weigert and M. Sára, *J. Membrane Sci.* 106, 147 (1995).
154. C. Weiner, M. Sára, and U. B. Sleytr, *Biotechnol. Bioeng.* 43, 321 (1994).
155. C. Weiner, M. Sára, G. Dasgupta, and U. B. Sleytr, *Biotechnol. Bioeng.* 44, 55 (1994).
156. Breitwieser et al., *BioTechniques* 21, 918 (1996).
157. U. B. Sleytr, H. Bayley, M. Sára, A. Breitwieser, S. Küpcü, C. Mader. S. Weigert, F. M. Unger, P. Messner, B. Jahn-Schmid, B. Schuste, D. Pum, K. Douglas, N. A. Clark, J. T. Moore, T. A. Winningham, S. Levy, I. Frithsen, J. Pankovc, P. Beale, H. P. Gillis, D. A. Choutov, and K. P. Martin *FEMS Microbiol. Rev.* 20, 151 (1997).
158. C. Nardin and W. Meier, *Rev. Mol. Biotechn.* 90, 17 (2002).
159. C. Nardin, M. Winterhalter, and W. Meier, *Langmuir* 16, 7708 (2000).
160. C. Nardin and W. Meier, *Chimia* 55, 142 (2001).
161. C. Nardin, T. Hirt, J. Leukel, and W. Meier, *Langmuir* 16, 1035 (2000).
162. H. S. Kim, J. D. Hartgerink, and M. R. Ghadiri, *J. Am. Chem. Soc.* 120, 4417 (1998).
163. K. Motesharei and M. R. Ghadiri, *J. Am. Chem. Soc.* 119, 1306 (1997).
164. A. B. Sieval, A. L. Demirel, J. W. M. Nissink, M. R. Linford, J. H. van der Maas, W. H. de Jeu, H. Zuilhof, and E. J. R. Sudhölter, *Langmuir* 14, 1759 (1998).
165. A. B. Sieval, R. Linke, H. Zuilhof, and E. J. R. Sudhölter, *Adv. Mater.* 12, 1457 (2000).
166. R. Kraayenhof, C. J. M. van Rijn, R. Kraayenhof, G. J. Sterk et al., *Biomembranes* 1284, 191 (1996).
167. B. A. Cornell, V. L. B. Braach-Maksvytis, L. G. King, P. D. J. Osman, B. Raguse, L. Wieczorek, and R. J. Pace, *Nature* 387, 580 (1997).
168. O. Braha, B. Walker, S. Cheley, J. J. Kasianowicz, L. Song, J. E. Gouaux, and H. Bayley, *Chem. Biol.* 4, 497 (1997).
169. G. Che, B. B. Lakshmi, C. R. Martin, E. R. Fischer, and R. S. Ruoff, *Chem. Mater.* 10, 260 (1998).
170. W.-H. Chu and M. Ferrari, *SPIE Proc.* 2593, 9 (JAAR).
171. J. L. Gole, R. P. Gao, Z. L. Wang, and J. D. Stout, *Adv. Mater.* 12, 1938 (2000).
172. M. Côté, M. L. Cohen, and D. J. Chadi, *Phys. Rev. B* 58, 4277 (1998).
173. R. Tenne, L. Margulis, M. Genut, and G. Hodes, *Nature* 360, 444 (1992).
174. S. Iijima, *Nature* 354, 56 (1991).
175. S. Iijima and T. Ichihashi, *Nature* 363, 603 (1993).
176. See, for example, P. M. Ajayan and T. W. Ebbesen, *Rep. Prog. Phys.* 60, 1025 (1997); M. S. Dresselhaus, G. Dresselhaus, and P. C. Eklund, "Carbon Nanotubes: Preparation and Properties." CRC Press, Boca Raton, FL, 1997.
177. A. Rubio, J. L. Corkill, and M. L. Cohen, *Phys. Rev. B* 49, 5081 (1994); X. Blase, A. Rubio, S. G. Louie, and M. L. Cohen, *Europhys. Lett.* 28, 335 (1994); A. Rubio, S. G. Louie, and M. L. Cohen, *Phys. Rev. B* 51, 6868 (1994).
178. Y. Miyamoto, A. Rubio, S. G. Louie, and M. L. Cohen, *Phys. Rev. B* 50, 360 (1994).
179. Y. Miyamoto, M. L. Cohen, and S. G. Louie, *Solid State Commun.* 102, 605 (1997).
180. N. G. Chopra, R. J. Luyken, K. Cherrey, V. H. Crespi, M. L. Cohen, S. G. Louie, and A. Zettl, *Science* 269, 966 (1995).
181. S. Iijima and T. Ichihashi, *Nature* 363, 603 (1993).
182. D. S. Bethune, C. H. Kiang, M. S. de Vries, G. Gorman, R. Savoy, J. Vazques, and R. Beyers, *Nature* 363, 605 (1993).
183. P. M. Ajayan, J. M. Lambert, P. Bernier, L. Barbedette, C. Colliex, and J. M. Planiex, *Chem. Phys. Lett.* 215, 50 (1993).
184. J. M. Lambert, P. M. Ajayan, P. Bernier, and J. M. Planiex, *Chem. Phys. Lett.* 226, 364 (1994).
185. C. Journet, W. K. Maser, P. Bernier, A. Loiseau, M. Lamy de la Chapelle, S. Lefrant, P. Deniard, R. Lee, and J. E. Fischer, *Nature* 388, 756 (1997).
186. H. Dai, A. G. Rinzler, P. Nikolaev, A. Thess, D. T. Colbert, and R. E. Smalley, *Chem. Phys. Lett.* 260, 471 (1996).
187. Z. F. Ren, Z. P. Huang, J. W. Xu, J. H. Wang, P. Bush, M. P. Siegal, and P. N. Provencio, *Science* 282, 1105 (1998).
188. V. I. Merkulov, D. H. Lowndes, Y. Y. Wei, G. Eres, and E. Voelkl, *Appl. Phys. Lett.* 76, 3555 (2000).

189. C. Bower, W. Zhu, S. Jin, and O. Zhou, *Appl. Phys. Lett.* 77, 830 (2000).
190. L. Delzeit, I. McAninch, B. A. Cruden, D. Hash, B. Chen, J. Han, and M. Meyyappan, *J. Appl. Phys.* 91, 6027 (2002).
191. K. B. K. Teo, M. Chhowalla, S. B. Lee, D. G. Hasko, H. Ahmed, G. A. J. Amaratunga, and W. I. Milne, University of Cambridge, UK.
192. M.-H. Hong, K. H. Kim, J. Bae, and W. Jhe, *Appl. Phys. Lett.* 77, 2604 (2000).
193. D. Ugarte, T. Stöckli, J. M. Bonard, A. Châtelain, and W. A. de Heer, *Appl. Phys. A* 67, 101 (1998).
194. A. W. Adamson and A. P. Gast, "Physical Chemistry of Surfaces," 6th ed. Wiley, New York, 1997.
195. E. Dujardin, T. W. Ebbesen, A. Krishnan, and M. M. J. Treacy, *Adv. Mater.* 10, 1472 (1998).
196. T. Werder, J. H. Walther, R. L. Jaffe, T. Halicioglu, F. Noca, and P. Koumoutsakos, *Nano Letters* 1, 687 (2001).
197. A. Gil, J. Colchero, M. Luna, J. Gómez-Herrero, and A. Baró, *Langmuir* 16, 5086 (2000).
198. J. Li, C. Papadopoulos, J. M. Xu, and M. Moskovits, *Appl. Phys. Lett.* 75, 367 (1999).
199. Anopore membranes, Whatman Ltd.
200. Aquamarijn Micro Filtration BV, www.microsieve.com.
201. H. Asoh, K. Nishio, M. Nakao, T. Tamamura, and H. Masuda, *J. Electrochem. Soc.* 148, 152 (2001).
202. C. R. Martin, M. Nishizawa, K. Jirage, M. Kang, and S. B. Lee, *Adv. Mater.* 13, 1351 (2001).
203. Y. Kobayashi and C. R. Martin, *Science* 268, 700 (1995).
204. S. Yu, S. B. Lee, M. Kang, and C. R. Martin, *Nano Letters* 1, 495 (2001).
205. S. B. Lee and C. R. Martin, *Chem. Mater.* 13, 3236 (2001).
206. S. B. Lee and C. R. Martin, *Anal. Chem.* 73, 768 (2001).
207. B. B. Lakshmi and C. R. Martin, *Nature* 388, 758 (1997).
208. Y. Kobayashi and C. R. Martin, *Anal. Chem.* 71, 3665 (1999).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanomotor F_1-ATPase

Pia D. Vogel

Southern Methodist University, Dallas, Texas, USA

CONTENTS

1. INTRODUCTION

It has been known for a long time that enzymes and proteins undergo large conformational changes while they perform their biological tasks. The catalytically active regions of enzymes need to sequentially open to bind their specific substrates; they then close to allow the specific chemical reaction to occur in a defined chemical environment and finally open up again up to allow the product to be released.

Other proteins have evolved to perform physical work by actively transporting or moving other molecules or proteins within cellular systems. Examples of these type of proteins are the following: (a) The kinesin motor proteins "walk" along "tubulin-rails" transporting vesicles along given pathways in the cell. (b) The sliding of actin protein filaments on myosin motor filaments cause the contraction of whole subcellular organizations in muscle cells during muscle contraction or allow migration of eukaryotic cells. (c) Other proteins such as dynein together with tubulin permit movement of whole, free cells like sperm cells and protozoa. Many of these systems require energy in the form of adenosine-5′-triphosphate (ATP) to power the respective conformational transitions. (d) Bacteria use the energy stored in proton gradients that are generated across a phospholipid membrane to drive the rotary movement of flagella. These are elongated structures that enable bacteria to swim to their source of nutrition.

All in all, coordinated movement is a concept that nature has used and perfected for billions of years. Only within the past few years, when humans have started to design smaller and smaller artificial devices, reaching into the nanometer scale, have these cellular and subcellular natural organizations been termed with names such as "nanomotors." However, the design of these motor proteins is intriguing, and it is fascinating to envision the potential use of biological devices within nanomechanical arrangements.

In this chapter I will focus on an enzyme whose physiological function is not that of a physical motor. The enzyme was designed to synthesize ATP, Nature's universal energy currency, by using the energy inherent in food sources or in the sunlight (depending on the origin of the enzyme). The enzyme is called the F_0F_1-ATP synthase, and it is, like ATP itself, ubiquitous in all organisms. The ATP synthase has intrigued researchers for more than four decades for a variety of different reasons. In an article entitled "The ATP Synthase—A Splendid Molecular Machine," published in 1997 [1], Paul D. Boyer from the Molecular Biology Institute at UCLA stated that "All enzymes are beautiful, but the ATP synthase is one of the most beautiful as well as the most unusual and important." Boyer shared the 1997 Nobel Prize for Chemistry with John E. Walker and Jens Skou. Both Boyer and Walker received the prize for their accomplishments in elucidating the structure and mechanism of the ATP synthase. In this chapter I will discuss how the ATP synthase accomplishes its natural goal, the synthesis of ATP, while also being Nature's smallest and so far most efficient molecular motor.

2. ATP SYNTHASE, STRUCTURE, AND MECHANISM

The F_0F_1-ATP synthase, an enzyme found in all organisms, has intrigued scientists for almost 40 years. This highly asymmetric protein assembly drives one of Nature's most challenging and probably the most important chemical reaction, namely the synthesis of ATP. The challenge of this reaction is the formation of an anhydride bond between adenosine-5′-diphosphate (ADP) and inorganic phosphate (P_i) in an aqueous environment that would usually favor ATP hydrolysis and not the synthesis reaction. The energy

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (83–89)

for the synthesis reaction is provided by the flow of protons down an electrochemical gradient and across the corresponding energy coupling membrane. In bacteria this membrane is the plasma membrane, in eukaryotes it is the mitochondrial membrane, and in photosynthetic organisms this process mostly takes place in the thylakoid membrane of chloroplasts. In mitochondria and bacteria, the synthase efficiently couples the oxidation of NADH to the formation of ATP. Nicotinamide adenine dinucleotide in the reduced form (NADH) is a redox equivalent that is generated when an organisms digests nutritious compounds like sugars, fats, or proteins. The reducing energy (in the form of two electrons) that is stored in NADH is utilized in the respiratory chain (Fig. 1), where the electrons are transferred in stepwise redox processes to oxygen. This process leads to the reduction of oxygen to form water while NADH is oxidized to NAD^+ (Fig. 1). During these redox processes protons (H^+) are pumped across the energy-coupling membranes by the membrane enzymes involved. The energy stored in the generated proton gradient drives the ATP synthesis as we will see next.

In photosynthetically active organisms, a light-driven proton translocation is coupled to the endergonic formation of the terminal phosphoryl anhydride bond of ATP ($ADP + P_i \rightarrow ATP$).

Under certain *in vivo* conditions, for example, in some bacteria when they live in an anaerobic environment, the reaction of the synthase can be reversed so that an ATP hydrolysis-driven proton pumping occurs.

Under physiological conditions, the free energy that is stored in the terminal (β–γ)-anhydride bond of an ATP molecule is about -9×10^{-20} J, equaling about -50 kJ/mol of ATP, depending on the cellular conditions.

The ATP synthases from different organisms are highly conserved. Much of the research on F_0F_1-ATP synthases has focused on the bacterial enzyme from *Escherichia coli*, mainly because molecular biological tools are easily available, but also because the *E. coli* enzyme is simpler in subunit composition and can be readily purified in rather large quantities. A second favored source for purifying the ATP synthase or substructures thereof is the thermophilic bacterium PS3. In addition to easy availability, the PS3 enzyme is also very stable at ambient and even elevated temperatures.

In all organisms the synthase consists of two components: the membrane embedded F_0 sector that is responsible for the translocation of protons and the membrane associated F_1 part that contains the sites for nucleoside triphosphate synthesis (Fig. 2). Both parts of the enzyme are highly asymmetric. The F_0 part of the *E. coli* enzyme consists of three different polypeptides a, b, and c with a stoichiometry of 1a, 2b, and 9–12 copies of the c subunit per functional unit [2]. Subunits a, b, and c are membrane integral and a and c are directly involved in proton movement through the membrane (for review, see [3]). Subunit b has an N-terminal membrane anchor. The residual amino acids of the protein protrude into the cytoplasm and interact, at least partially, with the F_1 part of the synthase, linking it to the membrane.

F_1 is the catalytic part of the synthase and consists of the five different subunits α through ε with three copies each of the major subunits α and β and one copy each of the minor subunits, γ, δ, and ε [4]. If this F_1 part is removed from the membrane and uncoupled from the proton gradient, it is only capable of net ATP hydrolysis, owing to the lack of driving force provided by the proton gradient. This part of the enzyme is therefore also named the F_1-ATPase.

It has been known since the 1980s that six nucleotide binding sites are located on the α and β subunits. The three binding sites that are located on the β subunits are

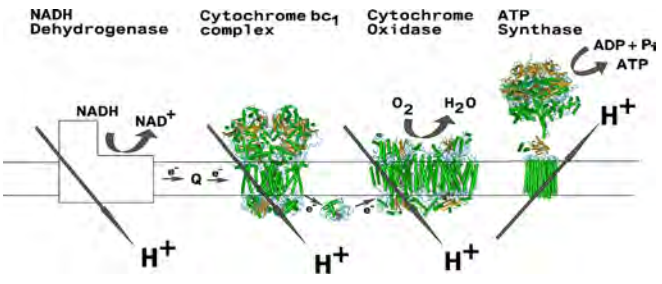


Figure 1. The respiratory chain. Electrons are transferred from NADH and H^+ to an enzyme called NADH dehydrogenase. From there the electrons are further transported along the membrane systems using the mobile carriers, coenzyme Q and cytochrome *c*. At the respective enzymes, NADH dehydrogenase, cytochrome bc_1 complex, and cytochrome *c* oxidase, the energy from the electron transfer is transduced into pumping of protons across the membrane. The resulting proton gradient drives the synthesis of ATP by the F_0F_1-ATP synthase. Where possible, the X-ray structural models of the corresponding proteins were created using Cn3D version 4.0 from the NCBI. The molecular coordinates of the protein structures can be found in the Protein Data Bank (cytochrome bc_1 complex, PDB-file: 1BGY; cytochrome *c*, PDB-file: 1GIW; cytochrome *c* oxidase, PDB-file: 1OCZ; and F_0F_1-ATP synthase, PDB: 1QO1).

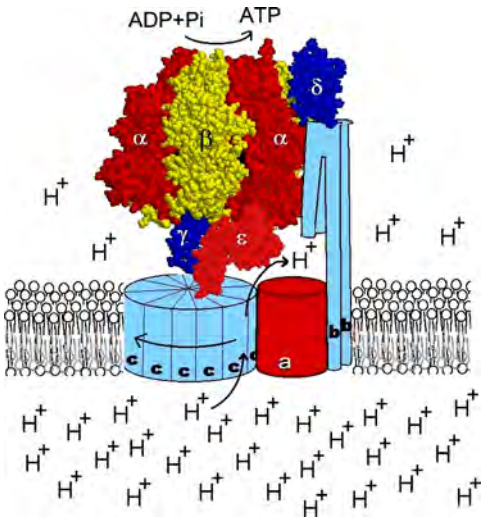


Figure 2. Model of the F_0F_1-ATP synthase. The diagram shows the two portions of the ATP synthase. The polypeptides a, b, and c are at least partially embedded in the membrane and are constituents of the F_0 portion of the membrane. Subunit a and the ring of c subunits are involved in the translocation of protons across the membrane. The membrane-associated F_1 part consists of subunits α, β, γ, δ, and ε and is responsible for the synthesis of ATP. The subunits $ab_2\alpha_3\beta_3\delta$ are considered to be the stator to the rotary motor. Subunits $c_{9-14}\gamma\varepsilon$ were shown in different experiments to rotate during catalysis (for more information see text).

catalytically active (catalytic sites) and show strong positive catalytic cooperativity. When all three sites are filled with substrate, the rate of catalysis is at least 10^5 times faster than if only one substrate is bound to the catalytic sites; for more details see [5–7]. The function of the three sites located on the α subunits remains unclear, and they are usually referred to as noncatalytic sites.

In the late 1970s, Boyer presented a model for the reaction mechanism of the ATPase that was termed the "binding change mechanism." One of the tenets of the theory stated that every one of the catalytically active nucleotide binding sites of the ATPase is in a different structural and functional state at any given time during catalysis and during sequential catalytic cycles; all of the sites go through the same states. While one of the sites is ready to bind the substrates, ADP and P_i, a second site has both these substrates bound, and a third site is in a conformation that allows the actual bond formation between ADP and P_i, resulting in elimination of water and the synthesis of ATP. The binding change mechanism model further states that the energy for the synthesis of ATP is solely required to release the product from the active site and not for the actual formation of the (β–γ)-anhydride bond. Experiments indicated that the equilibrium between ADP + P_i and ATP in the catalytic site is close to unity. The idea that these sites sequentially participate in this mechanism suggested that binding of substrate at one site promotes the release of product at a second site. This proposed mechanism beautifully accommodated the data that had been gathered about structure, subunit composition, and catalysis of the enzyme. One important question that remained at this point was what type of conformational transition enables the catalytic sites to sequentially change their conformation and function. The first suggestion that a rotation of internal subunits was part of the catalytic mechanism was published in the early 1980s, also by Boyer [8, 9] (for review see [10]).

Tremendous progress in our knowledge about the mechanism of ATP hydrolysis and synthesis was made in the past decade, gaining even more momentum when a high resolution structural model of the beef heart mitochondrial F_1-ATPase was presented in 1994 by Walker's group [11]. An X-ray crystallographic model of the beef heart mitochondrial ATPase is shown in Figure 3. The model gives detailed information about the structure of the nucleotide binding sites. In the model two of the catalytic sites are rather similar in conformation, whereas the third catalytic site differs significantly from the others in structure. This third side does not have a nucleotide bound (at least in the crystal structure), and a whole section of the binding site is rotated to form a very open conformation (able to release product). The model shows all the noncatalytic sites to be very similar in structure.

Several other structural models of various subassemblies of F_1 or F_0F_1 from a variety of different sources have become available over the last few years [12–16], increasing our knowledge about the structural basis of the ATPase and ATP synthase.

The structural models made rational mutagenesis of key residues within the nucleotide binding sites possible and allowed researchers to monitor the occupancy of the nucleotide sites during turnover using biophysical techniques

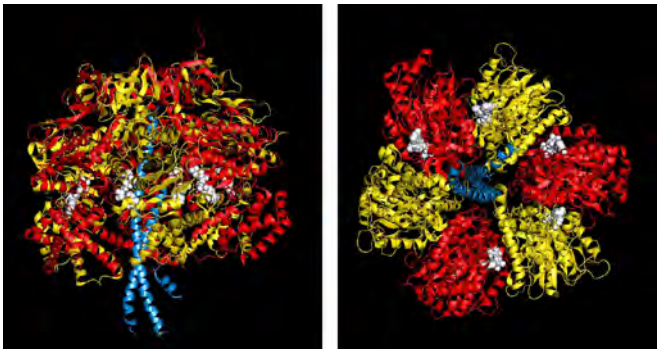


Figure 3. X-ray structural model of the mitochondrial F_1-ATPase [10] Left: side view, right: view from the membrane surface. The molecular coordinates of a variety of different F_1-ATPases from different sources are available in the Protein Data Bank. Here the model is based on the coordinates of the beef heart mitochondrial enzyme as described in [10]. The model was created using the program Rasmol and the molecular coordinates from the Protein Data Bank.

[17]. These biophysical data suggested that at maximal turnover conditions and saturating ATP, all three catalytic sites are occupied in F_1 as well as in solubilized F_0F_1 [17–20]. Experiments from our group using electron spin resonance spectroscopy and spin-labeled nucleotides suggested that under equilibrium conditions almost all nucleotide binding sites are filled [21–23].

3. ROTATIONAL MECHANISM OF THE F_1-ATPase

For about 15 years researchers tried to provide evidence that rotation of some of the subunits in F_1 was indeed a part of the mechanism of the ATPase. Especially after the X-ray structural models of the F_1 clearly placed the γ-subunit into the central cavity inside the ring of three α and three β subunits and showed its asymmetric interactions with the $\alpha_3\beta_3$ ring, it became likely that γ may be the subunit rotating relative to $\alpha_3\beta_3$. In different chemical and biophysical approaches, Cross's [24, 25] and Junge's [26] groups presented strong evidence that indeed such rotation was taking place. However it took until 1997 for Yoshida's group to unequivocally show in real time that F_1-ATPase is indeed a molecular motor [27]. In Figure 4 we present a diagram of the experimental set-up that Noji et al. [27] used to show that rotation occurs. A substructure of the F_1-ATP synthase from the thermophilic bacterium PS3 was used. This substructure only contains the subunits α and β as well as the γ subunit that resides in the middle of the $\alpha_3\beta_3$ ring. This assembly is very stable and hydrolyzes ATP efficiently. The β subunits of the construct were genetically modified to carry a series of histidine amino acid residues, each at one end of the polypeptide. These polyhistidines or "His-tags," as they are called, bind tightly to nitrilo triacetic acid that has a Ni^{2+} ion chelated to the carboxylic acid moieties. This compound is commonly abbreviated as Ni-NTA. In the "rotation" experiment, glass coverslips were covered with Ni-NTA. In a second step, the His-tagged F_1 was added to the top of the coverslips which resulted in F_1 that was immobilized "head down" onto the plates through

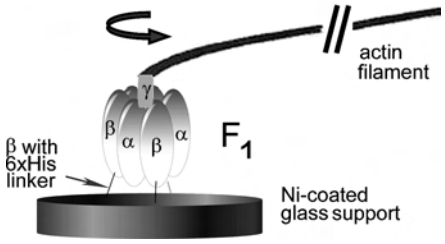


Figure 4. Diagram of a F_1-ATPase–actin construct as it was used to show rotation of the γ-subunit relative to $\alpha_3\beta_3$ (for details of the experiment see [27]). A substructure of the F_1-ATPase that consists only of $\alpha_3\beta_3\gamma$ from the thermophilic bacterium PS3 was used and genetically engineered to contain polyhistidine residues (His-tags) at the β subunits. The His-tags were tightly bound to glass coverslips that are covered with a Ni–NTA substrate (see text). The γ subunit was further genetically and chemically modified to carry biotin that again bound tightly to streptavidin, a protein with strong affinity to 4 molecules of biotin. Actin monomers were also chemically modified with biotin, polymerized under appropriate conditions, and finally decorated with actin-specific fluorescent dyes. The preformed actin filaments were attached to the F_1 motor via the streptavidin protein. Rotation of the actin filament was observed in a fluorescence microscope.

its His-tags. The γ subunit that protrudes from the other end of the $\alpha_3\beta_3\gamma$ substructure was also genetically modified to contain cysteine residues that can be reacted with sulfhydryl-specific reagents. Here the cysteines were modified with a bifunctional reagent that introduced a biotin molecule to the γ subunit of F_1. Biotin is a small organic molecule that very tightly and specifically binds to a protein called strepavidin. Streptavidin contains four binding sites for biotin, so after binding of streptavidin to the biotin molecule that was attached to the γ subunit, biotinylated actin filaments were added to the construct. We mentioned actin before as another example of a part of a molecular motor. Actin, as part of the contractile apparatus of the muscle and part of the cytoskeleton of individual cells, is a globular, monomeric protein that spontaneously polymerizes to form stable filaments under appropriate conditions. In the experiment described by Noji et al. actin monomers were biotinylated and subsequently allowed to polymerize. In addition, a fluorescent dye was introduced into the experiment, which made the formation of the filaments visible under appropriate excitation for fluorescence to occur. Addition of such modified actin filaments was the final step in the visualization of a very impressive molecular motor. With a fluorescence microscope, Noji and colleagues were able to observe the rotation of the γ subunit in real time.

In the initial experiments, they observed rotation of actin filaments of up to 4 μm in length in the presence of ATP for several minutes and hundreds of revolutions. The rotation always occurred counterclockwise when viewed from the direction of the membrane (or here from the top). Between 0.2 and 10 revolutions/s were observed, depending on the length of the filaments attached (1–4 μm). To put this into a "macroscopic" perspective, the "height" of the F_1-construct from the glass plate to the end of subunit γ is about 15 nm, and the length of the actin filament that can be rotated by this motor is up to 4 μm, or 4000 nm, which makes it about 260 times the size of the F_1 molecular motor. If one imagines F_1 to be the size of a human being, about 1.50 m tall, that motor would be able to rotate a construct that is 400 m long and roughly 1 m in diameter in an aqueous environment.

It was, however, not a trivial undertaking to observe the rotation. Thousands of assembled F_1-actin filament constructs had to be monitored and stringent controls were necessary to show without a doubt that ATP hydrolysis-dependent rotation of the γ subunit really took place. Considering the complicated setup of the experiment and the dimensions of the reaction partners, this is perhaps not so surprising indeed. In the meantime, a number of research groups have repeated the experiments using F_1-ATPases from different sources and of different substructure composition.

More recently, various research groups were able to show that the ring of c subunits of F_0 within the membrane rotates. When the driving force of the proton gradient across a membrane is strong enough, the downhill flow of the protons drives the rotation of the ring of c subunits within the membrane, resulting in the rotation of the γ subunit. It appears that the rotation of the γ subunit in itself causes the necessary conformational change to allow ATP to be released from the ATP synthase. In the opposite direction, hydrolysis of ATP resulted in the rotation of subunit γ as was shown in the soluble, nonmembrane-associated system. This subunit γ rotation then caused the rotation of the ring of subunits c, resulting in the pumping of protons across the membrane in an ATP-driven fashion [28–31] (for reviews, see [32 and 33]).

A diagram of the construct used to explore the rotation of the ring of c subunits is shown in Figure 5. In a recent paper [34] it was shown, using chemical cross-linking of various subunits of the F_0F_1 synthase, that when the rotor subunits c are immobilized onto Ni–NTA-coated glass plates using histidine tags and a Ni–NTA surface, the subunits of the synthase that usually do not participate in rotation (a, b_2, and the $\alpha_3\beta_3$ ring, also called the stator subunits) rotated when ATP was added as a fuel to the artificially prepared membrane systems. The results strongly indicated that the

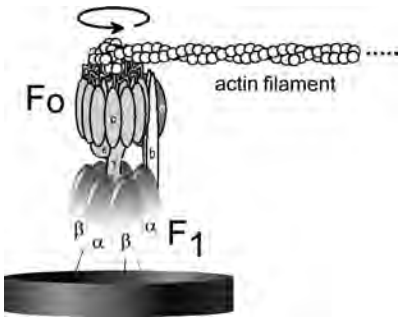


Figure 5. Diagram of a construct to show the rotation of the ring of subunits c (for more details see [30]). In addition to the histidine tag that allows the F_1 portion of the enzyme to be immobilized on glass slides, a second affinity tag (here a so-called Strep-tag) was introduced to the ring of subunits c. Strep-tags efficiently bind to streptavidin or related proteins and also bind to biotin molecules that were attached chemically to a fluorescently labeled actin filament. The rotation of the actin filament again can be observed in a fluorescence microscope.

membrane-embedded F_0F_1-ATP synthase could be divided into mechanical units that rotate relative to each other upon powering the system with fuel, ATP, or a H^+ gradient.

4. ENERGY AND TORQUE OF ROTATION

Isolated F_1-ATPase hydrolyzes ATP at a rate of about 300 molecules/s. For the load-free motor this would correspond to 100 revolutions/s because 3 ATP molecules are hydrolyzed per full rotation. The rotation of subunit γ was shown to be a three-step process [35–39], with each step corresponding to the hydrolysis of 1 ATP molecule. At low ATP concentrations (1 μM) even smaller rotational substeps of 90° and 30° were resolved [39] that were attributed to the binding of ATP and the release of ADP. Using the F_1-actin filament construct with 1 μm filament length and low ATP concentrations that only allow slow rotation, Yasuda et al. [40] observed distinct steps where the actin filament rested in 120° intervals.

Calculated from the rotational velocity, the torque reached during rotation is >40 pN·nm. By comparison, the sliding force that linear motors like myosin and kinesin (that were mentioned in the Introduction) generate is much smaller. Myosin produces a sliding force of 3–6 pN [41, 42] and kinesin a force of 5 pN [43]. The energy required to generate the torque in F_1 is about 8×10^{-20} J for a 120° rotation step. The free energy that is provided by the hydrolysis of one ATP (one ATP is hydrolyzed for a 120° rotation) is about 9×10^{-20} J. In other words, Nature has designed here a nanomotor, the ATP synthase, which runs at more than 90% efficiency. The efficiency may even be higher if the viscoelastic mechanics of the ATPase–actin construct are considered.

In two "head-to-head" papers [44, 45], Junge and coworkers studied the effects of slow viscoelastic relaxation of the actin filament on the torque generated by the F_0F_1 holoenzyme. The construct they used is depicted in Figure 6.

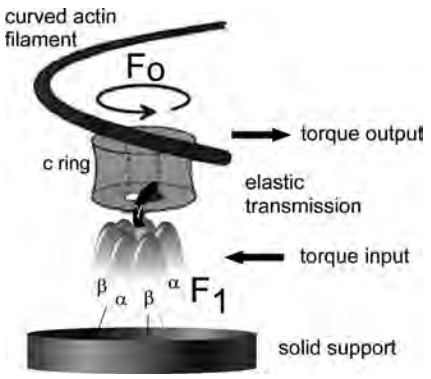


Figure 6. Diagram of an experiment to determine the torque generated by the ATPase (for more details see [44]). The basis of the experiments to determine the torque generated by the F_0F_1-ATP synthase is as described in the legend to Figure 3. The F_1 part is fixed to glass slides using oligo-histidine tags and a Ni–NTA-coated surface. The ring of c subunits carries Strep-tags that allow the actin filament to bind. The analysis of the viscoelasticity of the actin filament was performed using video microscopy in an inverted fluorescence microscope.

The results of these experiments and theoretical analysis suggest that indeed the torque that is generated by the synthase (here being powered by the electrochemical proton gradient) exceeds 40 pN · nm and was estimated to be more in the range of 50 ± 6 pN · nm. There were only very small variations in the torque output of the motor, implying a soft, elastic power transmission between the two parts of the motor, F_0 and F_1. This elastic transmission seems to be an essential feature to allow this biomotor to perform under such high turnover rates. The higher torque value that was calculated also shows that indeed this nanomotor works with almost perfect energy efficiency.

5. POWERING NANOELECTROMECHANICAL SYSTEMS

Very soon after details of the rotary mechanism became known, researchers started to investigate the nanomechanical world. Montemagno and his co-workers pioneered in this work and reported the "powering [of] an inorganic nanodevice with a biomolecular motor" in *Science* in 2000 [46, 47]. It is interesting to note that at the time this research was performed, Montemagno and his group were located at the Cornell University in Ithaca, NY. From this same university some 40 years before, Racker, Penefsky, and co-workers reported that they had isolated a soluble "factor" from beef heart mitochondria that hydrolyzed ATP [48]. They reported that this factor could restore ATP synthesis in mitochondrial membranes that had lost this ability during isolation. They called this interesting protein the "factor 1" or F_1 for ATP synthesis.

The Montemagno group used a substructure of the F_1-ATPase from the thermophilic bacterium PS3 that contains only the basic rotor assembly consisting of the subunits α, β, and γ, similar to the experiments described by the Yoshida group [27]. In initial experiments, procedures for the specific attachment and positioning of the nanoelectromechanical systems (NEMS) were tested [47]. Using electron beam lithography, an array pattern was etched on a 25 mm coverslip. The coverslips were then patterned with metal substrates such as gold, copper, or nickel. A synthetic peptide containing a polyhistidine tag as described above was allowed to attach to the metal-coated coverslips. The peptide had a microsphere covalently attached to the opposite side, and laser tweezers were used to test the strength of the attachment to the metal surface. Nickel was found to be the most promising candidate, and, therefore, in their following paper [46], the group presented a method to generate nanofabricated Ni posts of a diameter of 50–120 nm and a height of 200 nm. Ni posts were chosen as a "stand" for binding the ATPase motor to prevent problems that are associated with "dragging" the "propeller" unit in close proximity to the base of the motor. Nickel propeller rods were generated and their dimensions were optimized for optical detection and minimal friction during the rotation process. The propellers were then coated with biotinylated peptides that also carried histidine tags. The polyhistidines allowed binding of the peptides to the Ni rods, while the biotinylation allowed

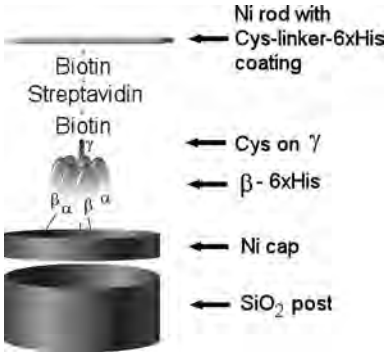


Figure 7. Diagram of a nanoelectromechanical system as developed by the Montemagno group, (for more details see [46]). An F_1 substructure containing $\alpha_3\beta_3\gamma$ is immobilized via His-tags onto Ni post structures with a diameter of 50–120 nm and a height of 200 nm. Subunit γ was biotinylated and using the biotin–streptavidin–biotin construct, a spacer is formed that binds to prefabricated Ni rods that have been covered with His-tagged oligopeptides. Upon addition of ATP, the ATP-driven rotation of the γ subunit can be visualize as a rotation of the Ni rod.

the interaction of the propellers with F_1 that carried a streptavidin protein at the γ subunit. Functional NEMS were constructed by sequential addition of the individual components: F_1-ATPase was biotinylated at its γ subunit through specific modification of engineered cysteine residues in the appropriate position. The biotinylated F_1 was placed onto the Ni posts by virtue of the histidine tags that were located at the β subunits of the constructs. Next, streptavidin was bound to the biotin molecules attached to the γ subunit and finally the Ni propellers decorated with biotinylated peptides were added and assembled onto the motor. One model for an appropriate design is shown in Figure 7.

Finally, and most ingeniously, the F_1 part of the NEMS was further modified to contain a chemical switch that allowed the system to be turned on or off, without having to run the motor out of its power source, ATP [49]. Using computational design methods [50–53] and molecular biological methods Hellinga and co-workers constructed a binding site for Zn^{2+} ions. Using the original construct of the motor with an actin filament attached to the γ subunit, they showed that rotation occurred in the absence of Zn^{2+} ions, was inhibited in the presence of Zn^{2+}, and could be reestablished when Zn^{2+} was removed from the mixture. A very similar approach was used by Hisabori and co-workers [54] who introduced a chemical switch into the γ subunit that could be turned off in the presence of oxidants such as $CuCl_2$ and switched back on upon reduction with dithiothreitol, a reducing agent, and thioredoxin, which is found to be a natural regulator of the ATPase that is found in chloroplasts.

6. OUTLOOK

The first steps for the ATP synthase into the "nano-world" have already been successfully performed. Whether it will be used as a pump in nanoscale chemical reaction chambers, a three-step switch to modulate flow of fluids or electricity, or even a motor to drive nanosubmarine structures, it will be solely up to the imagination of researchers, both in engineering and in natural sciences, to see the future and limitations of use for such highly intriguing devices.

GLOSSARY

Adenosine-5′-diphosphate (ADP) Corresponds to ATP with one less phosphoryl group.

Adenosine-5′-triphosphate (ATP) A biological molecule that is used in all organisms for short-term energy storage.

Biotinylation Chemical linking of biotin to another biomolecule.

His-Tag A series of (usually) six histidine residues that are genetically attached to recombinantly produced proteins. The histidine residues form complexes with Ni^{2+} ions and allow affinity purification of the corresponding proteins.

Inorganic phosphate (P_i) Salt of phosphoric acid.

Micromolar (μm) One micromole of a compound dissolved in 1 liter of a solvent. One micromole equals 10^{-6} mole.

Nanoelectromechanical system (NEMS) An inorganic particle (here a Ni rod) that is linked to a biological system (here the F_1-ATPase).

Ni-Nitrilo triacetic acid (NTA) Nitrilo triacetic acid that has Ni^{2+} ions chelated to the carboxylic acid moieties. The reagent is commonly used for affinity chromatography to specifically bind proteins that have oligo histidine tags attached.

ACKNOWLEDGMENTS

The author thanks Dr. John G. Wise for critically reading the manuscript and preparing the artwork.

REFERENCES

1. P. D. Boyer, *Annu. Rev. Biochem.* 66, 717 (1997).
2. D. L. Foster and R. H. Fillingame, *J. Biol. Chem.* 257, 2009 (1982).
3. G. Deckers-Hebestreit and K. Altendorf, *Annu. Rev. Microbiol.* 50, 791 (1996).
4. J. E. Walker, I. M. Fearnley, N. J. Gay, B. W. Gibson, F. D. Northrop, S. J. Powell, M. J. Runswick, M. Saraste, and V. L. Tybulewicz, *J. Mol. Biol.* 184, 677 (1985).
5. R. L. Cross and C. M. Nalin, *J. Biol. Chem.* 257, 2874 (1982).
6. J. G. Wise, T. M. Duncan, L. R. Latchney, D. N. Cox, and A. E. Senior, *Biochem. J.* 215, 343 (1983).
7. C. Kayalar, J. Rosing, and P. D. Boyer, *J. Biol. Chem.* 252, 2486 (1977).
8. P. D. Boyer and W. E. Kohlbrenner, in "Energy Coupling in Photosynthesis" (B. Selman and S. Selman-Reiner, Eds.), p. 231. Elsevier Science, New York, 1981.
9. P. D. Boyer, in "Biochemistry of Metabolic Processes" (D. L. Lennon, F. W. Stratmann, and R. N. Zahlten, Eds.), p. 465. Elsevier Biomedical, New York, 1983.
10. P. D. Boyer, *Biochim. Biophys. Acta* 1140, 215 (1993).
11. J. P. Abrahams, A. G. Leslie, R. Lutter, and J. E. Walker, *Nature* (*London*) 370, 621 (1994).
12. M. J. van Raaij, J. P. Abrahams, A. G. Leslie, and J. E. Walker, *Proc. Natl. Acad. Sci. U.S.A.* 93, 6913 (1996).
13. J. P. Abrahams, S. K. Buchanan, M. J. Van Raaij, I. M. Fearnley, A. G. Leslie, and J. E. Walker, *Proc. Natl. Acad. Sci. U.S.A.* 93, 9420 (1996).

14. Y. Shirakihara, A. G. Leslie, J. P. Abrahams, J. E. Walker, T. Ueda, Y. Sekimoto, M. Kambara, K. Saika, Y. Kagawa, and M. Yoshida, *Structure* 5, 825 (1997).
15. D. Stock, A. G. Leslie, and J. E. Walker, *Science* (*Washington, DC*) 286, 1700 (1999).
16. M. A. Bianchet, J. Jullihen, P. L. Pedersen, and L. M. Amzel, *Proc. Natl. Acad. Sci. U.S.A.* 95, 11065 (1998).
17. J. Weber, S. Wilke-Mounts, R. S. Lee, E. Grell, and A. E. Senior, *J. Biol. Chem.* 268, 20126 (1993).
18. J. Weber and A. E. Senior, *Biochim. Biophys. Acta* 1319, 19 (1997).
19. J. Weber, S. D. Dunn, and A. E. Senior, *J. Biol. Chem.* 274, 19124 (1999).
20. S. Löbau, J. Weber, and A. E. Senior, *Biochemistry* 37, 10846 (1998).
21. S. Burgard, J. H. Nett, H. E. Sauer, Y. Kagawa, H. J. Schafer, J. G. Wise, P. D. Vogel, and W. E. Trommer, *J. Biol. Chem.* 269, 17815 (1994).
22. R. M. Lösel, A. H. Erbse, J. H. Nett, J. G. Wise, G. Berger, G. Girault, and P. D Vogel, *Spectrosc. Acta A* 52, 73 (1996).
23. R. M. Lösel, J. G. Wise, and P. D. Vogel, *Biochemistry* 36, 1188 (1997).
24. T. M. Duncan, V. V. Bulygin, Y. Zhou, M. L. Hutcheon, and R. L. Cross, *Proc. Natl. Acad. Sci. U.S.A.* 92, 10964 (1995).
25. Y. Zhou, T. M. Duncan, V. V. Bulygin, M. L. Hutcheon, and R. L. Cross, *Biochim. Biophys. Acta* 1275, 96 (1996).
26. D. Sabbert, S. Engelbrecht, and W. Junge, *Nature* (*London*) 381, 623 (1996).
27. H. Noji, R. Yasuda, M. Yoshida, and K. Kinosita, Jr., *Nature* (*London*) 386, 299 (1997).
28. Y. Sambongi, Y. Iko, M. Tanabe, H. Omote, A. Iwamoto-Kihara, I. Ueda, T. Yanagida, Y. Wada, and M. Futai, *Science* (*Washington, DC*) 286, 1722 (1999).
29. T. Suzuki, H. Ueno, N. Mitome, J. Suzuki, and M. Yoshida, *J. Biol. Chem.* 227, 13281 (2002).
30. O. Pänke, K. Gumbiowski, W. Junge, and S. Engelbrecht, *FEBS Lett.* 472, 34 (2000).
31. S. P. Tsunoda, R. Aggeler, M. Yoshida, and R. A. Capaldi, *Proc. Natl. Acad. Sci. U.S.A.* 98, 898 (2001).
32. R. L. Cross, *Biochim. Biophys. Acta* 1458, 270 (2000).
33. M. Yoshida, E. Muneyuki, and T. Hisabori, *Nat. Rev.* 2, 669 (2001).
34. K. Nishio, A. Iwamoto-Kihara, A. Yamamoto, Y. Wada, and M. Futai, *Proc. Natl. Acad. Sci. U.S.A.* 99, 13448 (2002).
35. D. Sabbert and W. Junge, *Proc. Natl. Acad. Sci. U.S.A.* 94, 2312 (1997).
36. D. Sabbert, S. Engelbrecht, and W. Junge, *Proc. Natl. Acad. Sci. U.S.A.* 94, 4401 (1997).
37. K. Hässler, S. Engerlbrecht, and W. Junge, *FEBS Lett.* 426, 301 (1998).
38. K. Adachi, H. Yasuda, H. Noji, M. Harada, M. Yoshida, and K. Kinosata, Jr., *Proc. Natl. Acad. Sci. U.S.A.* 97, 7243 (2000).
39. R. Yasuda, H. Nuji, M. Yoshida, K. Kinosita, Jr., and H. Itoh, *Nature* (*London*) 410, 898 (2001).
40. R. Yasuda, H. Noji, K. Kinosita, Jr., and M. Yoshida, *Cell* 93, 1117 (1998).
41. J. T. Finer, R. M. Simmons, and J. A. Spudich, *Nature* (*London*) 368, 113 (1994).
42. H. Miyata, H. Yoshikawa, H. Hakozaki, N. Suzuki, T. Furuno, A. Ikegami, K. Kinosita, Jr., T. Nishizaka, and S. Ishiwata, *Biophys. J.* 68, 286 (1995).
43. K. Svoboda, C. F. Schmidt, B. J. Schnapp, and S. M. Block, *Nature* (*London*) 365, 721 (1993).
44. O. Pänke, D. A. Cherepanov, K. Gumbiowski, S. Engelbrecht, and W. Junge, *Biophys. J.* 81, 1220 (2001).
45. D. A. Cherepanov and W. Junge, *Biophys. J.* 81, 1234 (2001).
46. R. K. Soong, G. D. Bachand, H. P. Neves, A. G. Olkhovets, H. G. Craighead, and C. D. Montemagno, *Science* (*Washington, DC*) 290, 155 (2000).
47. C. Montemagno and G. Bachand, *Nanotechnology* 10, 225 (1999).
48. M. E. Pullman, H. S. Penefsky, A. Datta, and E. Racker, *J. Biol. Chem.* 235, 3322 (1960).
49. J. J. Schmidt, X. Jiang, and C. D. Montemagno, *Nano Lett.*, in press.
50. H. Liu, J. J. Schmidt, G. D. Bachand, S. S. Rizk, L. L. Looger, H. W. Hellinga, and C. D. Montemagno, *Nat. Mat.* 1, 173 (2002).
51. H. W. Hellinga and F. M. Richards, *J. Mol. Biol.* 222, 763 (1991).
52. J. S. Marvin and H. W. Hellinga, *Proc. Natl. Acad. Sci. U.S.A.* 98, 4955 (2001).
53. D. Benson, A. E. Haddy, and H. W. Hellinga, *Biochemistry* 41, 3262 (2002).
54. D. Bald, H. Noji, M. Yoshida, Y. Hirono-Hara, and T. Hisabori, *J. Biol. Chem.* 276, 39505 (2001).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanoparticle Drug Delivery to the Brain

K. Ringe, C. M. Walz, B. A. Sabel

Otto-von-Guericke Universität Magdeburg, Magdeburg, Germany and Nanopharm AG, Center for Neuroscience Innovation and Technology, Magdeburg, Germany

CONTENTS

1. INTRODUCTION

The blood–brain barrier (BBB) protects the brain against toxic substances that circulate in the bloodstream. Although this is life-supporting protection for the brain, the existence of the BBB is a severe limitation for the delivery of most drugs to the brain because they do not cross the BBB in sufficient amounts. A large number of potentially useful drugs, such as cytostatics and central nervous system (CNS)-active agents, do not cross the BBB at all or in insufficient quantities. However, for therapeutic reasons methods to increase the bioavailability of drugs in the brain are needed to deliver drugs to the brain which are usually blocked from entering the brain by the BBB.

One possible way to achieve this goal is to attach drugs to nanoparticles and thus transport them across the BBB. Nanoparticles are solid colloidal particles, ranging in size from 1 to 1000 nm (usually 200–300 nm), and they are a rather useful "drug delivery system" to target drugs to the brain. Here, we review the structure and the role of the BBB and describe the manner whereby molecules normally pass through the BBB. Furthermore, we will discuss how nanoparticles can be prepared and purified and review their physicochemical properties and drug release mechanisms. We then discuss the evidence that nanoparticles can be used to deliver drugs to the brain in the living animal and document their usefulness in some animal models.

2. TRANSPORT OF MOLECULES ACROSS THE BLOOD–BRAIN BARRIER

In some cases potentially useful compounds to treat brain disorders are injected into the bloodstream or given orally, but they do not reach the brain at all or not in sufficient amounts (insufficient bioavailability). Therefore, therapeutic efficiency to treat brain diseases is diminished or prevented because systemic administration of the drug does not lead to an effective brain concentration [1–4]. There are many reasons why this may be the case: the molecules may be too large, they have unfavorable physiochemical properties (such as polar functional groups), they may be metabolized by enzymes before reaching their target (the brain), or they may be extruded at the cerebrovascular endothelium well before reaching the brain cells (neurons) upon which they should act. The major reason why drugs do not reach the brain is the existence of the BBB [5–7]. To develop methods to overcome this barrier, a good understanding of its nature is required.

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (91–104)

2.1. The Physiology of the Blood–Brain Barrier

The BBB is not one single structure or membrane in the brain, but it is created by the way the blood vessels in the brain are organized. Thus, understanding the BBB requires an understanding of the anatomy and physiology of the blood vessels in the brain. Both large and small capillaries form a richly branched and complex network throughout the entire brain tissue. Like a chimney made of individual bricks, the brain blood vessels consist of a monolayer of endothelial cells that are connected with each other by tight junctions (zonulae occludentes) [8]. The part of the cell's membrane facing the bloodstream is called the "luminal" membrane [9–11], and the side which is exposed to the actual brain tissue is called the "abluminal" membrane. This part faces the extracellular liquid of the brain parenchyma where pericytes and endfeet of astrocytes surround the blood vessels.

The most important site of the BBB lies at the cerebral microvessels, that is, the very fine vessels that have extremely small diameters. Because endothelial cells are very polarized, that is, essentially similar to the epithelium, they exhibit very low pinocytic activity and possess a high number of mitochondria that are needed for the multiple energy-dependent active transport mechanisms found in endothelial cells [12].

Peripheral vessels in the rest of the body can much more easily transport molecules across their membrane because they are fenestrated and have many active transcellular transport mechanisms. In contrast, in central blood vessels of the brain, even small molecules like antibiotics have great difficulty crossing the barrier and only a limited number of molecules can actively cross the endothelial cells. Here, the endothelial cells use specific transport systems to allow the influx of glucose, iron, amino acids, peptides, small organic acids, and others. This is necessary so that substances which are critical for brain metabolism and function can gain fast and efficient access to the brain via specific energy-dependent carrier mechanisms at the endothelium [13].

2.2. The Role of Pericytes

Pericytes, which are located on the abluminal side of the endothelial cells, are also part of the BBB. Pericytes are a physiological heterogeneous cell population and are found on all microvessels in nearly every organ [13], but they never cover the entire blood vessel. The pericytes, which are located on the "brain side," are encapsulated by the basal membrane of the endothelial cells, and they are responsible for the synthesis and release of different components of the basal membrane and the extracellular matrix such as collagen and glycosaminoglycocan [14]. The basal membrane of the endothelial cells and that of the pericytes are closely attached to each other so that both cells have a common basal membrane [15]. Electron microscopy studies have revealed fenestrations between pericytes and endothelial cells [16], and pericytes have contractile properties that may play a role in the regulation of the blood flow. Several molecules are involved in these contractile functions such as actin, myosin, tropomyosin, vimentin, and desmin [17–19]. Pericytes have an important role in the function of the BBB. They are responsible for the maintenance of the barrier function and the stability of the vessel [20].

2.3. The Role of Astrocytes

Astrocytes (which are brain glia cells) also contribute to the BBB, and they are attached with their endfeet to the pericytes and the endothelial cells. Astrocytes are glial cells responsible for the homeostasis and the ion regulation in the brain [20], but their endfeet cover the blood vessels only partially. In contrast to endothelial cells and pericytes, astrocytes are not connected to other cells by tight junctions, and they do not have a common basal membrane. Therefore, polar molecules (such as proteins) can enter the interstitial liquid and be directly transported to the pericytes and the endothelial cells [21].

That astrocytes are important for the induction and maintenance of the BBB properties can be deduced from the following observations in cell cultures: in the presence of astrocytes or medium conditioned by astrocytes, endothelial cells express markers important for BBB characteristics and develop tight junctions [9–11]. On the other hand, endothelial cells promote the development and differentiation of astrocytes. This interaction between both cell types actually occurs even when there is no contact between the two cell types, indicating that some soluble, extracellular factors are mediators of BBB development.

3. TRANSPORT MECHANISMS OF THE BLOOD–BRAIN BARRIER

Although the BBB is a fairly "tight" structure when "unfriendly" molecules try to enter the brain, it does have several mechanisms to allow "friendly" molecules to get in. Among them are active, carrier-mediated transport mechanisms for relatively small molecules, an absorptive-mediated endocytosis mechanism for positively charged peptides, and receptor-mediated endocytosis mechanisms specific to certain peptides. For the structure of the BBB see Figure 1.

3.1. Transport of Anionic Compounds

Various anionic compounds are transported across the BBB by the monocarboxylic acid transporter (MCT) [22, 23]. Lactic acid, for example, is transported by a specialized mechanism and not by nonionic passive diffusion [24, 25].

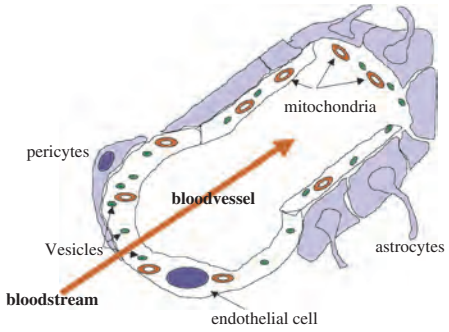


Figure 1. The blood brain barrier (BBB).

MCT1 is present in the BBB and acts as a transport mechanism for lactic acid and other monocarboxylic compounds [26].

3.2. Transport of Cationic Compounds

Cationic compounds enter the brain by passive diffusion and also by carrier-mediated transport, which is believed to be the mechanism of action whereby several drugs cross the BBB [27–32]. An endogenous hydrophilic amine, choline, has been demonstrated to be taken up by a carrier-mediated transport mechanism [29–31], and a carrier-mediated amine transport system is likely to be responsible for transport of cationic compounds across the BBB.

3.3. Transport of Peptides

Because peptides are usually relatively large, hydrophilic, and unstable, efficient permeation into the brain generally does not occur [33]. However, certain relatively small peptides are transported by carrier-mediated transport mechanisms, and others cross the BBB by transcytosis including receptor-mediated transcytosis and absorptive-mediated transcytosis.

3.4. Carrier-Mediated Transport of Peptides

Because the transport of glutathione across the BBB is saturable [34, 35], carrier-mediated transport that has a high efficiency comparable to the brain uptake of single amino acids such as phenylalanine and cysteine is probably involved.

3.5. Absorptive-Mediated Endocytosis

Absorptive-mediated transcytosis is triggered by electrostatic interactions between the positively charged moiety of the peptide and the negatively charged plasma membrane surface region [36–39]. If one attempts specific targeting to the CNS, absorptive-mediated transcytosis is not expected to be useful, because absorptive-mediated transcytosis or endocytosis also functions in other tissues such as kidney and liver [38].

3.6. Receptor-Mediated Endocytosis

To get peptides to cross the BBB they need to be manipulated in a manner such that penetration into the brain is achieved. Sometimes this is achieved by synthesis of chimeric peptides. They are formed by covalent binding of the nonpermeable but pharmacologically effective portion of the peptide to an appropriate vector that can be transported across the BBB [38]. In this case, the chimeric peptide is first transported into the brain endothelial cytoplasm by receptor-mediated or absorptive-mediated endocytosis. The intact chimeric peptide is then transferred into the brain's interstitial space by receptor-mediated exocytosis. Subsequently, the binding between the vector and the pharmacologically active peptide is cleaved and, finally, the released peptide exerts its pharmacological effect in the brain [36, 38, 40, 41].

3.7. Efflux by P-Glycoprotein

Several anticancer drugs such as vinca alkaloids and anthracyclines exhibit lower accumulation in the brain than would be expected from their lipophilic properties [42, 43]. Furthermore, tumor cells show a resistance to anticancer drugs such as anthracyclines and other nonrelated compounds [44]. This multidrug resistance is accompanied by overexpression of the transmembrane P-glycoprotein and a decrease of drug concentration in the resistant cells compared with that in drug-sensitive cell lines [45, 46]. P-glycoprotein molecules are located at the luminal side of the endothelial cells and their function is to serve as an energy-dependent efflux pump and transport various drugs out of cells, thus decreasing their accumulation in the cytoplasm, and, as a consequence, reducing their efficiency [46–48]. P-glycoprotein has a very broad substrate specificity [49–51].

3.8. Transport of Drugs across the BBB

Given the highly selective nature of the BBB, the question arises as to how drugs that usually do not cross the BBB can still be delivered to the brain. An estimated 99% of all potential drugs are clinically useless because the BBB prevents them from getting into the brain. Thus, finding ways to overcome the BBB are of great clinical significance.

There are two fundamental ways to get drugs into the brain. The first and most invasive method is application of the drug directly into the brain. This can be done by intracerebral injection or infusion or by artificial drug delivery systems such as polymer implants that release drugs in a sustained manner [52]. For example, anticancer drugs are often applied by intraventricular infusion, but this is a rather invasive technique with high risks. In view of the rapid turnover of the cerebrospinal fluid, however, the pharmacological compound is rapidly removed from the brain into peripheral blood vessels and the diffusion-dependent concentration in the parenchyma is far below desirable levels.

The second and most favorable method is to manipulate or "trick" the BBB, so that neurologically active compounds can be given systemically by oral, intramuscular, or intravenous application.

Here, several different approaches are possible: osmotic opening of the BBB or pharmacologically by raising the lipophilicity of the drug using chimeric molecules or by attaching drugs to nanoparticles [53].

3.9. Osmotic Opening of the BBB

The BBB can be opened by osmosis or by biochemical and pharmacological means [54–56]. Application of hyperosmotic mannitol and arabinose solutions results in the local reversible destruction of tight junctions, thus increasing the permeability of the endothelial cells. For example, hyperosmolar urea solution withdraws water from the endothelial cells and leads to shrinkage of the cells and consequently to an opening of the tight junctions. As one can easily imagine, the major disadvantage of this method is that the brain loses its protection from neurotoxic substances for a certain period of time, which results in significant unwanted side effects such as brain edema [54–56].

A similar approach is the infusion of a vasoactive substance such as bradykinin or leukotriene, which increases vesicular transportation activity of the endothelial cells.

3.10. Pharmacological Approaches

Among the pharmacological approaches the most common one is increasing the lipophilicity of the drug and using chimeric proteins as well as cationized antibodies. Chimeric proteins can be transported via receptor-mediated or adsorptive-mediated transcytosis [57, 58]. Prodrugs are an excellent example of such drug manipulations. With this method, the original compound is manipulated to make it more lipid-soluble, providing greater brain penetration. However, increased lipid solubility may significantly alter pharmacokinetic parameters such as clearance and half-life, which, for chlorambucil derivatives, is undesirable [58].

The third and in our mind the most desirable method to get a drug across the BBB is the novel technique of nanoparticle drug delivery. Compounds can be attached to or incorporated into nanoparticles, which serve as a universal "drug delivery system," allowing many pharmacological agents to cross biological barriers.

4. NANOPARTICLES AS A DRUG DELIVERY TOOL FOR BRAIN TARGETING

Over the last few years work in the laboratories of Kreuter [59, 60] and Sabel [61–64] has demonstrated the usefulness of nanoparticles as a universal tool to deliver drugs to the brain. The evidence is consistently mounting that important diseases of the brain can be treated by a combined nanoparticle/drug approach, and brain tumor treatment, in particular, has been accomplished in animal models. A variety of experiments have been carried out to date using various animal species and different behavioral paradigms and disease models, all of which substantiate the value of nanoparticles as a novel drug delivery method.

To understand the usefulness and potential of nanoparticle technology for drug delivery to the brain, we shall first describe methods of preparation, the mechanism of drug binding and release, and evidence of toxicity and efficacy. We will also briefly discuss practical matters that are important for clinical use of nanoparticles such as sterilization procedures, stability, and purification issues and then discuss the evidence indicating their usefulness as a tool for delivery of drugs to the brain.

5. PREPARATION OF NANOPARTICLES

There are different methods to manufacture nanoparticles: (1) emulsion polymerization, (2) interfacial polymerization, (3) solvent evaporation, (4) solvent deposition, and (5) denaturation. The properties of nanoparticles vary with different polymers, stabilizers, and surfactants used during the manufacturing process. Each excipient added may have an influence on the bioavailability of the drug carried, the brain drug uptake, and the stability of the drug in the plasma [65].

5.1. Emulsion Polymerization

Emulsion polymerization is one of the most rapid and most frequently used methods for nanoparticle preparation. Here, the monomer is added to a continuous phase, usually an aqueous phase at room temperature under constant stirring conditions. It is also possible to use an organic phase as the continuous phase. The polymerization can be initiated either by free radicals or by ion formation. The polymerization is initiated by the reaction of a monomer molecule with an initiator molecule. Triggers for the initiation of the reaction can be ultraviolet (UV) light, hydroxyl ions, or high-energy radiation. The polymer chain starts to grow when these initiated monomer ions or monomer radicals react with other monomer molecules. Additional monomer is solubilized in surfactant micelles or emulsified in larger droplets. After completion of polymerization, the reaction mixture is filtered, neutralized, and purified by centrifugation to remove any residual monomer. An example for an anionic process is the preparation of "poly(alkylcyanoacrylate) nanoparticles" and an example for a free radical-initiated emulsion polymerization is the manufacturing of polymethyl methacrylate nanoparticles [66]. The process of emulsion polymerization has numerous advantages. Compared with other methods, it is rapid and in general there is no need to use stabilizers and surfactants. In addition, for industrial requirements it is easily scaled up [67]. In contrast, with the requirement of UV light, radiation, or free radicals to initiate the polymerization process, the incorporation of proteins and peptides during the polymerization is not possible. Furthermore, if one wishes to scale up this procedure, the need for purification of the nanoparticles via dialysis and centrifugation represents a problem [68]. The mechanism of an anionic polymerization is shown in Figure 2.

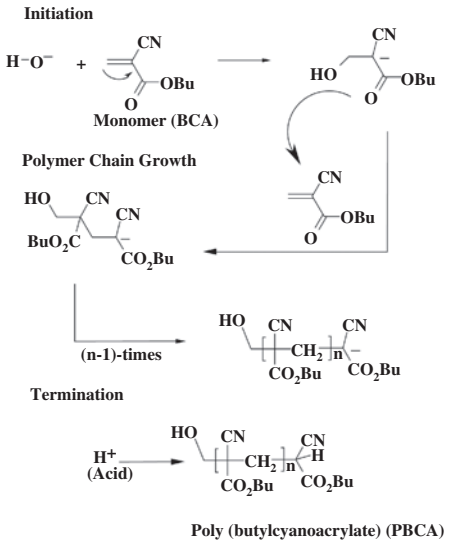


Figure 2. Mechanism of polymerization.

5.2. Interfacial Polymerization

To achieve interfacial polymerization, monomers are polymerized at the interface between two immiscible phases. Interfacial polymerization takes place in a medium consisting of an aqueous and an organic phase, which are homogenized, emulsified, or micro-fluidized by vigorously mechanical stirring. Al Kouhri Fallouh et al. [69] introduced the formation of polyalkylcyanoacrylate nanocapsules. In this process the monomer and the drug are dissolved in a mixture of oil and ethanol (oil/ethanol: $\frac{1}{10}$ to $\frac{1}{200}$) and then slowly added through a small tube or needle to an aqueous phase containing surfactants (poloxamer 188 or 407 or phospholipids). The oil used can be Miglyol or benzylic acid. The primary disadvantage of this method is the occurrence of strong shear forces. This excludes the possibility of adding proteins and peptides during the polymerization process for incorporation purposes. The monomer spontaneously polymerizes and forms nanocapsules that consist of an oil droplet and a polymeric shell [69]. An advantage of this process is that the drug is encapsulated into the nanocapsule and not just adsorbed onto the surface. This would protect it from enzymes, thus preventing premature biodegradation before it reaches the blood brain barrier.

5.3. Solvent Evaporation

The solvent evaporation method is a well-established and frequently used method for the manufacturing of particles with sizes above 1 μm and also sizes of less than 1000 nm. In this process the preformed polymer and the drug are dissolved in a volatile, water-immiscible organic solvent. This organic phase is then added to the aqueous phase under stirring, and the organic solvent is removed by heating and/or under reduced pressure. The polymer precipitates and forms micro- or nanospheres instantaneously, containing the drug dispersed in the polymer matrix network. The particles are then purified by filtration and centrifugation [70]. Examples of this process are the manufacturing of poly(lactic acid) nanoparticles and poly(lactic-coglycolic acid) nanoparticles [71, 72].

5.4. Solvent Deposition

In this process the polymer, e.g., poly(DL-lactide), and phospholipids are dissolved in a volatile organic solvent such as acetone. Then a solution of the drug in benzyl benzoate is added to the organic phase and the reaction mixture is poured into the water phase, which contains poloxamer 188 under moderate stirring conditions. Nanocapsules consisting of an oily core and a poly(lactic acid) shell are formed instantaneously. The organic solvent is then removed under reduced pressure. Partial removal of water also occurs [73, 74].

5.5. Denaturation

Nanoparticles can also be produced by denaturation of natural macromolecules such as albumin and gelatin in an oil emulsion. For this procedure the macromolecule is dissolved in an aqueous solution and the drug is entrapped in an oil emulsion. Then both phases are emulsified using a homogenizer. The size of the forming particles depends on the stirring velocity, slit width, and power of the homogenizer. Then the particles are hardened by cross-linking with an aldehyde, by heat denaturation, or by cooling below the gelation point [75–77].

6. PURIFICATION

All processes mentioned above require purification except for the emulsion polymerization process in some cases, when no surfactant and no organic solvents are used to produce nanoparticles. In general, the organic solvent is removed under reduced pressure and the resulting particles are purified by ultracentrifugation, ultrafiltration, gel chromatography, dialysis, or a combination of these methods. The choice of the appropriate purification method depends on the release properties of the nanoparticles. Inappropriate purification methods can lead to the loss of biologically active agent. Stabilization of the nanoparticles is usually achieved by lyophilization using a cryoprotector (such as mannitol) to prevent agglomerations [66].

7. STABILITY

The stability of poly(butylcyanoacrylate) nanoparticles was examined in water, in phosphate-buffered saline, in acidic medium (0.01 or 0.1 N HCl), and in human blood serum [78, 79]. The authors found that in acidic medium no formation of agglomerates occurred and no degradation took place. The nanoparticles could therefore be stored in acidic medium for several months. In contrast, when suspended in water or phosphate-buffered saline, the nanoparticles agglomerate. Nanoparticle polymers did not degrade in solution as indicated by the fact that they kept their size constant. In human blood serum no formation of agglomerates was observed, and nanoparticles were stable for 8 days. This result supports the possibilty of the use of nanoparticles for intraveneous administration, because microembolisms due to agglomeration are not expected.

8. STERILIZATION

An important prerequisite for the use of nanoparticles clinically is sterilization of the nanoparticle/drug complex. The possibility of sterilizing poly(butylcyanoacrylate) nanoparticles was studied by Sabel et al. [80]. Nanoparticle suspensions and nanoparticle powders were submitted to autoclaving (121° C for 20 min with or without cooling to 70° C afterward) or treated with formaldehyde (60° C). The nanoparticles were prepared using different stabilizers, such as dextran 70,000, poloxamer 188, and polysorbate 85. In most cases the authors detected a significant increase in particle size. For lyophilized nanoparticles no powder was suitable for preparing injectable suspensions. In particular, the resuspension of nanoparticle powders caused problems, which could not be overcome by intensified sonication. Under certain conditions poloxamer 188 suspensions and polysorbate 80 suspensions led to injectable nanoparticle

suspensions. In summary, sterilization of nanoparticles after fabrication is possible under certain conditions, but the favored way to obtain aseptic material seems to be the production of nanoparticles under aseptic conditions.

9. CHARACTERIZATION

The most important feature of nanoparticles is their small size. Therefore, some understanding is needed of how to manipulate and measure the size of nanoparticles. There are many different parameters that influence the size of particles during the manufacturing procedure. These are the pH of the reaction medium, the stirring velocity, the slitwidth, and the duration of ultrasonification. Besides size, there are also other physical parameters important for nanoparticles such as density, molecular weight, crystallinity, release, and degradation properties. The surface charge, hydrophilicity, and hydrophobicity significantly influence the body distribution of nanoparticles and, respectively, the bioavailability of the bound drug. Kreuter [66] reviewed the most common physicochemical methods for the characterization of nanoparticles. The size of nanoparticles is determined using photon-correlation spectrometry. This method, which quantifies light scattering, determines the hydrodynamic diameter of the nanoparticles via Brownian motion. It is important to verify the results obtained by this method using electron microscopy, because larger particles, such as dust or accidential microbial contamination, could lead to incorrect results [81, 82].

10. DRUG LOADING

Drugs can be loaded onto nanoparticles by adding them to a solution that contains previously prepared nanoparticles or by adding them to the reaction mixture during the polymerization process. Both methods can supply (1) a solid solution of the drug in the polymer [83, 84], (2) solid colloidal nanoparticles with dispersion of the drug in the polymer [85], (3) adsorption of the drug onto the surface of the nanoparticle [86], (4) chemical binding of the drug to the polymer [87], or (5) no binding or incorporation at all. Figures 3 and 4 show the two possible ways of drug loading.

The amount of bound drug and the type of interaction of drug and nanoparticles depend on the chemical structure of the drug and the polymer and the conditions of drug loading. The determination of the adsorption isotherm is one

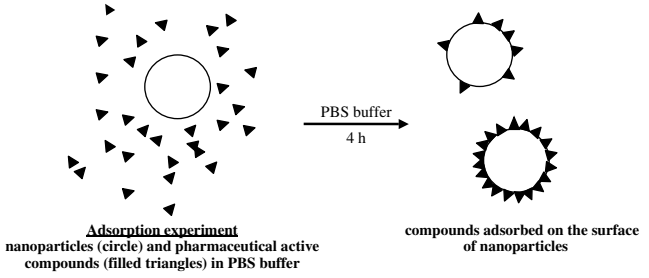


Figure 3. Drug loading of nanoparticles by adsorbtion.

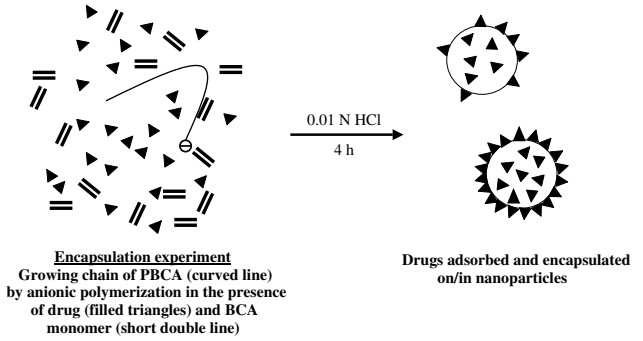


Figure 4. Drug loading of nanoparticles by encapsulation.

possible way to detect the type of binding and the binding rate (mg drug/mg nanoparticle) [88]. Linear sorption isotherms characterize solid solutions [84] and Langmuir- or S-type isotherms characterize surface adsorption [89]. Because nanoparticles are colloidal systems, precise determination of the drug content can be a problem. Therefore, the most reliable way is to separate the nanoparticles from the solution containing unbound drug by ultracentrifugation or gel filtration [90]. The amount of bound drug can be determined by subtracting the drug content in the supernatant from the primary amount of drug present in the suspension [66].

11. DRUG RELEASE

For manipulation of the rate and the timing of the drug release from nanoparticles, a good understanding of the mechanisms of drug release is needed.

There are five possible methods for drug release: (a) desorption of drug bound to the surface, (b) diffusion through the nanoparticle matrix, (c) diffusion through the polymer wall of nanocapsules, (d) nanoparticle matrix erosion, or (e) a combined erosion–diffusion process [66]. The pharmacokinetic analysis of drug release from nanoparticles can be described by a biexponential function

$$C_t = Ae^{-\alpha t} + Be^{-\beta t}$$

where C is the concentration of drug remaining in the nanoparticles at time t, A and B are system characteristic constants, and α and β are rate constants that can be obtained from semilogarithmic plots.

Gupta et al. [91] analyzed the mathematical problems of drug release using doxorubicin as the model drug. They proposed two main pathways of drug release. First, there is a rapid release of drug from the nanoparticles, which could stem from drug adsorbed on the nanoparticle surface. Then a slower, more controlled release takes place, which is possibly related to nanoparticle degradation or diffusion of the drug through the nanoparticle shell. Not only are the diffusion coefficient and the biodegradation rate main factors in drug release, but the biological environment also is important. Plasma proteins could be adsorbed onto the surface of the nanoparticles, so that an additional diffusion barrier is formed. On the other hand, nanoparticles could bind to

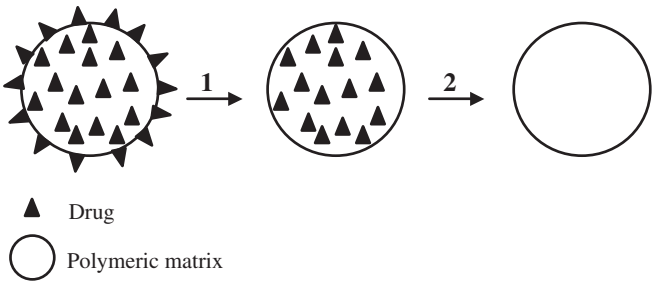


Figure 5. (1) Initial rapid desoption of drug from the nanoparticle surface; (2) Controlled release of drug by diffusion through the nanoparticle matrix and polymer wall. The release is dependent on the nanoparticle biodegradation of the polymer shell.

biological membranes, so that the transport of drugs through these membranes is facilitated. Hence, the results obtained by *in vivo* and *in vitro* drug release experiments are very different. Nevertheless, *in vitro* drug release experiments are very important for quality control reasons; see Figures 5 and 6 [66, 92, 93].

11.1. Ideal Properties of Polymeric-Based Nanoparticles

Before nanoparticles can be considered as a realistic drug delivery method, there are several properties the system needs to fulfill [94]. These are the following:

- stable in blood
- nontoxic
- nonthrombogenic
- nonimmunogenic
- noninflammatory
- no activation of neutrophils
- biodegradable
- avoidance of the reticuloendothelial system
- applicable to various molecules, such as small molecules, proteins, peptides, or nucleic acids (platform technology)
- scalable and inexpensive manufacturing process.

Because poly(butylcyanoacrylate) nanoparticles possess these properties they were reported to be useful tools for drug delivery to the brain in the several *in vivo* experiments. In addition, they exhibit very low (if any) toxicity and are considered to be relatively nontoxic. Their toxicity toward hepatocytes is rather low with an LD50 at about 0.4 mg/2 $\times$ 10^6 cells [95]. The LD50 in mice is 230 mg/kg after intraveneous application [96].

12. BODY DISTRIBUTION OF NANOPARTICLES—BRAIN TARGETING

Before discussing the evidence that nanoparticles are indeed useful to deliver drugs to the brain, we need to consider their fate after being injected into the body.

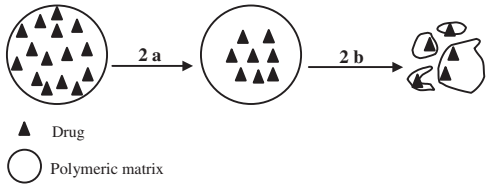


Figure 6. Combined nanoparticle erosion diffusion process. 2a: Diffusion of drug through the polymeric matrix and shell; 2b: Drug release through the biodegradation of the nanoparticle.

12.1. Fate of Nanoparticles in the Body

After intraveneous injection, nanoparticles are taken up very rapidly by the reticuloendothelial system (RES) and distribute especially into the liver (60–90%) and spleen (2–10%) and, to a minor degree, into the bone marrow. The RES consists of phagocytic cells originating from the bone marrow. The cells exist in the whole body, but their highest concentration is found in the liver (Kupffer cells), spleen, and bone marrow. Despite their great usefulness to target the brain, because of their high uptake by the RES after intraveneous injection, only a limited proportion of the nanoparticles actually reach the brain [97, 98]. Unless their surface is modified by special coatings (described below), nanoparticles themselves are of little value for delivery of drugs to the brain.

12.2. Poly(butylcyanoacrylate) Nanoparticles to Permit Transfer across the BBB

In recent years Kreuter et al. [59, 60] and Schroeder et al. [61–64] applied a special procedure to permit brain targeting of nanoparticles. This was accomplished by altering their surface with surfactants. Specifically, polymeric nanoparticles are coated with different hydrophilic surfactants. So far only poly(butylcyanoacrylate) nanoparticles possess the potential to transport drugs across the blood–brain barrier *in vivo* [99].

The most common method is to first fabricate the nanoparticles; then the drug is bound by adsorption on the surface of the nanoparticles or by forming a solid solution or dispersion. Thereafter, a surfactant is added to the nanoparticle/drug complex.

Poly(butylcyanoacrylate) nanoparticles have an average diameter size of about 200–300 nm and are very rapidly biodegraded [97, 100]. Figure 7 shows the two pathways of biodegradation of these nanoparticles. The main pathway consists of enzymatic cleavage of the butylester group of the polymer. The polymeric acid is then formed, and butanol is formed as well. Both metabolites are water soluble and will be excreted by the kidneys. During the other minor pathway, the polymer chain is degraded and formaldehyde is formed in traces, too low to be of physiological concern [101].

The first drug that was successfully delivered to the brain with this approach was the hexapeptide enkephalin, dalargin (Tyr-D-Ala-Gly-Phe-Leu-Arg). These studies were carried out first in the laboratory of Kreuter and subsequently and independently in the laboratory of Sabel.

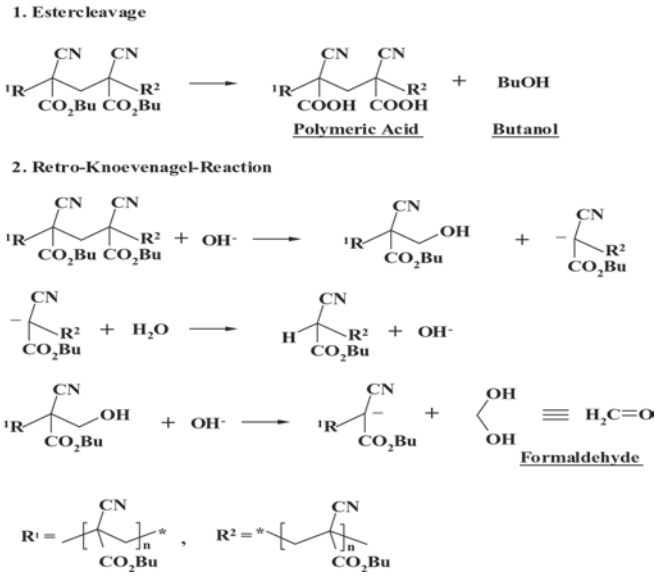


Figure 7. Biodegradation of PBCA-nanoparticles.

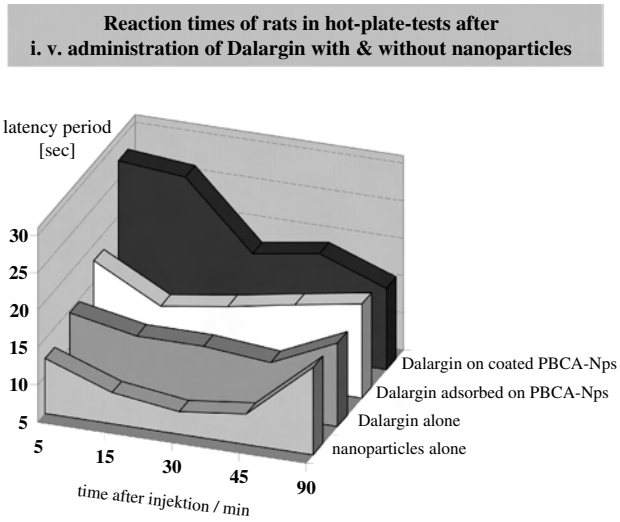


Figure 8. Central analgesic effect of dalargin-loaded nanoparticles coated with polysorbate 80 after i.v. injection.

12.3. Behavioral Evidence That Nanoparticles Can Deliver Drugs across the BBB Using Dalargin

Dalargin is an enkephalin-type peptide that has analgesic (pain-reducing) effects when injected directly into the brain. However, when given intravenously, it does not cross the BBB and no analgesia occurs [59, 102]. Studies to investigate the targeting of dalargin to the brain were carried out using the tail-flick test [59, 102] as well as the hot-plate test [61, 103]. Both tests are useful to quantify pain perception as a behavioral model of drug activity in the brain. When the rat tail is lightly pinched, the rat withdraws the tail by flicking it to the side. If an effective analgesic is given, the flicking does not occur because the rat feels no pain. Thus, one can quantify experimentally analgesic effects of a drug. In the hot-plate test the situation is similar: when placed on a hot plate, the animals will lick their paws. Time to tail flicking and time to paw withdrawal are very sensitive and reliable measures of pain perception and can thus be used to quantify analgesia.

To evaluate the usefulness of nanoparticles crossing the BBB, dalargin was bound to nanoparticles and injected into animals, which were then tested for analgesia. To be sure that drug alone or nanoparticles alone had no analgesic effect, groups of animals that received different treatments were compared: (1) dalargin alone; (2) polysorbate 80 alone; (3) poly(butylcyanoacrylate) nanoparticles alone; (4) a mixture of dalargin and polysorbate 80; (5) a mixture of dalargin and poly(butylcyanoacrylate); (6) a mixture of all three components, dalargin, polysorbate 80, and nanoparticles, mixed immediately before injection; and (7) a suspension of dalargin bound to nanoparticles without coating. After intravenous (i.v.) injection into mice, none of the groups showed an analgesic effect. Only dalargin bound to poly(butylcyanoacrylate) nanoparticles coated with polysorbate 80 produced a significant antinociceptive effect [59, 61, 102, 103]. This clearly showed that dalargin had crossed the BBB. Figure 8 shows the antinociceptive effect after i.v. administration of polysorbate 80-coated nanoparticles loaded with dalargin using the hot-plate test [61]. Apparently the coating with polysorbate 80 was the critical step in achieving the effect. This finding raised the question whether any other coating material may be useful as well.

Therefore, different coating materials in combination with the nanoparticles were used to study the specificity of the coating effect to deliver dalargin to the brain. It was found that polysorbate 20, 40, and 60 coatings have similar, but weaker, analgesic effects compared with polysorbate 80 coating. Poloxamers and poloxamines as well as surfactants of the Cremophor and Brij series [polyoxyethylene-(23)-laurylether] had no effect [60]. Thus, with a particular set of coating materials nanoparticles effectively delivered drugs to the brain after i.v. injection.

12.4. Oral Application of Peptides using Nanoparticle Technology

In the clinical setting it would clearly be desirable to avoid having to inject nanoparticles but rather to make them available also as an oral method for drug delivery. The efficacy of drug delivery to the brain using dalargin-loaded poly(butylcyanoacrylate) nanoparticles after oral application was studied by Schroeder et al. [62]. Different nanoparticle formulations containing dextran 12,000, poloxamer 188, or polysorbate 85 as stabilizer loaded with dalargin were used. No additional coating was used. The formulations were given intraveneously and orally. When polysorbate 85 was used as a stabilizer without coating, a significant analgesic effect occurred after i.v. injection as well as after oral application. Figure 9 shows the analgesic effect after oral application of polysorbate 85-stabilized nanoparticles loaded with dalargin. These results suggest that nanoparticles can also be used as a new tool for oral drug delivery of peptides [62].

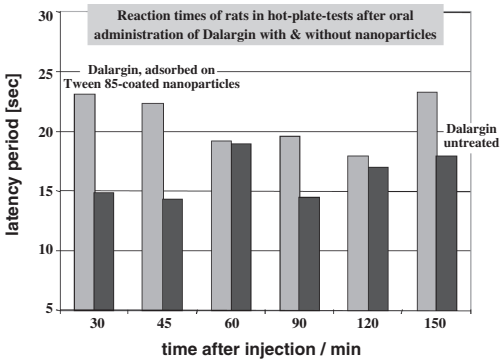


Figure 9. Central analgesic effect of polysorbate 85-stabilized nanoparticles loaded with dalargin after oral administration.

12.5. Nanoparticles as a Universal Method for Drug Delivery Across the BBB

Next the question arose whether such effects can only be achieved with the test drug dalargin or whether nanoparticles are a universal tool for delivery of drugs to the brain. Therefore, different antinociceptive agents, such as kytorphin [63] and loperamide [104] that normally cannot penetrate the BBB, were also tested using nanoparticles. Similarly to the dalargin experiments, analgesic effects of these drugs were observed. Drug solutions and uncoated nanoparticles, which served as controls, had no significant effect on analgesia, but a significant analgesic effect was obtained when nanoparticles coated with polysorbate 80 and polysorbate 85-stabilized nanoparticles without coating were used [63]. Pharmacokinetic studies of amitriptyline, a tricyclic antidepressant, bound to polysorbate 80-coated poly(cyanoacrylate) nanoparticles showed an enhancement in brain area under the curve (AUC) and a reduced serum AUC after i.v. injection. Similar results were also obtained with polysorbate 85-stabilized nanoparticles without coating [63]. AUC is the area under the curve of concentration versus time after adsorption of 20.0 mg/kg amitriptyline onto the surface of the nanoparticles after i.v. injection of amitriptyline-loaded nanoparticles in mice. The amitriptyline concentrations in serum and brain were determined after 5, 20, and 60 min after i.v. injection.

To better understand the mechanisms of action of the injected nanoparticle/drug complex, it is necessary to know its fate after injection. To this end the bioavailability and body distribution of radioactively labeled [^{3}H]-dalargin bound to poly(butylcyanoacrylate) nanoparticles coated with polysorbate 80 were studied by Schroeder et al. [64]. The studies were performed by comparing unbound ^{3}H-labeled dalargin with bound dalargin after i.v. injection into mice. The level of radioactivity in brain preparations was three times higher after i.v. injections of ^{3}H-labeled dalargin bound to poly(butylcyanoacrylate) nanoparticles and coated with polysorbate 80 after 5 and 20 min. In addition, radioactivity levels in the liver were decreased [64]. These results are compatible with the known time course of the analgesic effect found in behavioral studies 5 and 15 min after oral versus i.v. application [61, 62]. The increased dalargin level in the brain and the enhanced serum levels led to the conclusion that the drug is stabilized by the nanoparticles. This finding is also supported by the decreased liver levels. The nanoparticle formulation thus protects the labeled peptide from rapid degradation and uptake by the RES [64], thus permitting higher brain concentrations.

The question also arises whether these findings hold true only for analgesia studies or whether they generalize to other disease models as well. Here, a study by Alyautdin et al. [105] is of interest. They provoked the development of epileptiform seizures using tubocurarine bound to polysorbate 80-coated poly(butylcyanoacrylate) nanoparticles. Tubocurarine, a quaternary ammonium salt that is unable to penetrate the BBB, provokes epileptiform seizures upon direct injection into the brain. An *in-situ* rat brain perfusion technique was used and the electroencephalogram (EEG) was recorded. Tubocurarine-loaded polysorbate 80-coated nanoparticles caused the appearance of EEG seizures after 15 min. None of the controls, such as a tubocurarine solution, tubocurarine-loaded nanoparticles without polysorbate 80 coating, or a mixture of polysorbate and tubocuraine, affected on the EEG. In conclusion, only polysorbate 80-coated nanoparticles were able to transport tubocurarine across the BBB.

12.6. Influence of Nanoparticles on the Metabolism of Drugs

The use of poly(butylcyanoacrylate) nanoparticles as a transport system of antiepileptic drugs, such as valproic acid, was investigated by Schroeder et al. [106]. The authors evaluated the increase in the brain-to-serum ratio of the drug to reduce dose-related side effects in the periphery by using nanoparticles. Changes in the metabolism of valproic acid were also investigated. Mice were given four different i.v. injections: (1) valproic acid alone, (2) valproic acid + nanoparticles (dextran-stabilized), (3) valproic acid + nanoparticles (dextran-stabilized) + Tween 80, and (4) valproic acid + nanoparticles (polysorbate 85-stabilized). Twelve mice were decapitated 5, 20, and 60 min after i.v. injection. The blood was allowed to clot, and serum was taken after centrifugation. The brains were removed, cleaned, and frozen until analysis. The brain and serum levels of valproic acid and its metabolites were determined using a modified gas chromatography–mass spectroscopy assay [107]. It was found that the serum kinetics and the brain tissue levels of valproic acid did not change by administration of valproic acid/nanoparticle formulations compared with the drug alone. However, nanoparticles did influence the metabolism of valproic acid. They inhibited the metabolic degradation of the drug by mitochondrial β-oxidation but did not influence any other metabolic pathway. It was concluded that nanoparticles reduce the toxic side effects of valproic acid therapy by selectively blocking a metabolic pathway. Thus, nanoparticles may serve as a tool to change drug metabolism, probably through a sustained-release mechanism, in a desirable manner.

12.7. Nanoparticles as a Method to Delivery Cytostatic Agents to Treat Brain Tumors

In another line of research, drugs to treat brain tumors were investigated. Gulyaev et al. [108] bound doxorubicin to nanoparticles and then examined their body distribution after i.v. injection to rats. Rats were given four different formulations of doxorubicin (DOX): (1) DOX in saline, (2) DOX in polysorbate 80 in saline, (3) DOX bound to poly(butylcyanoacrylate) nanoparticles, and (4) DOX bound to poly(butylcyanoacrylate) nanoparticles overcoated with polysorbate 80. The two doxorubicin solutions showed a similar body distribution and no significant difference between plasma and organ (brain, liver, spleen, heart, or kidney) concentration. For both nanoparticle preparations a significant decrease in the heart concentration was observed. This is a very important finding, because the use of doxorubicin in the treatment of cancer is limited by its heart toxicity, which had also been noted by Couvreur et al. [109]. Only with the polysorbate 80-coated nanoparticle formulation was a high concentration of doxorubicin (>6 μg/g) achieved in brain and plasma. The low plasma level and brain concentrations (0.1 μg/g, below the detection limit) of the free drug without nanoparticles are caused by their rapid uptake by the RES [110–112]. In addition, polysorbate coating reduced the concentration of doxorubicin in the liver, spleen, and lung [108]. When taken together, these results suggest again that polysorbate 80-coated nanoparticles are a useful drug delivery vehicle for brain targeting. This and other studies by Gelperina et al. [113] described below open new possibilities for the treatment of brain cancer.

To evaluate the efficacy of a DOX–nanoparticle formulation, Gelperina et al. [113] investigated the treatment of rat brain tumors using doxorubicin-loaded, polysorbate 80-coated poly(butylcyanoacrylate) nanoparticles. To test efficacy, rats with intracranially transplanted glioblastoma 101/8 cells were treated with a DOX–nanoparticles formulation in the dose of 3×1.5 mg/kg. After 6 months these animals were killed and a beneficial effect was documented by histological examination of the brain. In the untreated control group all animals died within 17 days after implantation of cancer cells into the brain. In the remaining four groups [(1) nanoparticles + polysorbate (Ps), (2) DOX in saline, (3) DOX + Ps, and (4) DOX + nanoparticles] only one group (DOX + Ps) showed a long-surviving animal. This could have been due to the fact that the BBB might have been partially disrupted as a result of tumor growth. In conclusion, this study provides another indication that polysorbate 80-coated nanoparticles are a useful tool for drug delivery to the brain.

12.8. Other Approaches with Particles for Drug Delivery to the Brain

Nanoparticles are not the only method to increase drug bioavailability. Another approach to reduce RES uptake and to overcome the BBB is the preparation of magnetic microspheres combined with the application of a magnetic field to the brain of rats after intracarotid injection [114]. Two different types of microspheres of 1–2 μm diameter size were prepared: magnetic neutral dextran and cationic aminodextran microspheres. Both kinds of microspheres are prepared by an emulsification process using polysorbate 80 as the emulsifier. The magnetic particles were then injected into the carotid arteries of healthy rats and rats with implanted brain tumor (gliom-2/RG-2). A magnetic field was then applied to the brain. The rats were killed after 30 min and 6 h, and the magnetic particle concentration in the tissues was determined by atomic absorption spectroscopy. In healthy rats only low brain concentrations but high lung and spleen concentrations were found for both microspheres. In tumor-bearing rats, the brain concentration of magnetic cationic particles was increased compared with that of neutral particles. In general, these experiments with magnetic microspheres resulted in increased brain concentrations and decreased concentrations in peripheral organs, but did not show any significance because of the high variability among the groups.

13. MECHANISM OF NANOPARTICLE TRANSPORT ACROSS THE BLOOD–BRAIN BARRIER

Different mechanisms are described to explain the nanoparticle-mediated drug transport to the brain, including creation of a concentration gradient, a general surfactant effect, opening of the tight junctions, endocytosis, and inhibition of the efflux system [99]. Among them the most likely mechanism is the receptor-mediated endocytosis of nanoparticles. Lück et al. [115] found that apolipoprotein E (apo E) adsorbs on the surface of polysorbate 20-, 40-, 60-, or 80-coated nanoparticles after 5 min in human plasma at 37 °C. After separation of the particles from the serum by centrifugation, the adsorbed plasma proteins were desorbed and analyzed by two-dimensional gel electrophoresis. For polysorbate 20-, 40-, 60-, or 80-coated nanoparticles, apo E was found in the serum. In contrast, without coating or with overcoating of nanoparticles using poloxamers 338 and 407, Cremophor EL, or Cremophor RH40, no adsorption of apo E was observed. These results correspond to the findings that only polysorbate 20-, 40-, 60-, or 80-coated nanoparticles induced an antinociceptive effect using dalargin [60]. Kreuter et al. [116] studied the involvement of apo B and E in the mechanism of nanoparticle transport across the BBB. For i.v. injection into mice, poly(butylcyanoacrylate) nanoparticles loaded with dalargin were coated with apolipoproteins AII, B, CII, E, or J without or after precoating with polysorbate 80. In addition, nanoparticles loaded with loperamide were coated with apo E alone or after precoating with polysorbate 80. Furthermore, the antinociceptive effect of polysorbate-coated dalargin-loaded nanoparticles was measured in ApoEtm1Unc and C57BL/6J mice. The authors used the tail-flick test to determine the antinociceptive threshold. It was found that dalargin and loperamide-loaded nanoparticles coated with polysorbate 80 and/or with apo B or E induced the analgesic effect. Nanoparticles with polysorbate-precoating and apo B or E overcoating achieved a significantly higher analgesic effect. In addition, with the apo E-deficient ApoEtm1Unc mice, the effect was notably reduced compared with that in

C57BL/6J mice, from which the ApoEtm1Unc mice were derived. The results led to the conclusion that apo B and E are involved in the mediation of the transport of drugs bound to poly(butylcyanoacrylate) nanoparticles across the BBB. These results and also the studies of Lück et. al. [115] indicate that after i.v. injection apo E and/or B is anchored by the polysorbate on the surface of the nanoparticles [116]. Apo B and E are known to bind to lipoprotein receptors on the surface of cells [117, 118]. Low-density lipoprotein receptors have been identified in rat and monkey brains [119]. They also exist in the brain capillary endothelial cells [119, 120]. Polysorbate 80-coated nanoparticles adsorb the apolipoproteins after i.v. injection and thus seem to mimic lipoprotein particles that are able to interact with members of the low-density lipoprotein receptor family and are taken up via receptor-mediated endocytosis. Bound drugs may be further transported into the brain by diffusion [116].

14. CONCLUSION

Many drugs affecting the brain or spinal cord either do not cross the BBB at all or do not cross it in pharmacologically effective amounts. As a consequence, some drugs affecting the brain have undesirable peripheral side effects that pose a clinical problem. The severe disturbance caused by cytostatic agents used for the treatment of brain tumors are a dramatic example of this point. It is therefore desirable to have a method available whereby drugs can be delivered to the brain and spinal cord more effectively.

Nanoparticle technology is a very valuable method to deliver drugs across the BBB: nanoparticles have a diameter of about 200 nm and they are suspended in an aqueous solution. Depending on the method of the polymerization, drugs are either attached to the surface and/or are incorporated into poly(butylcyanoacrylate) particles. Subsequently, this drug–particle complex is covered by a suitable surfactant (such as polysorbate 80). This complex is then injected intravenously or given orally.

Nanoparticles can be stored either in lyophilized form or in solution. In lyophilized form they are stable for at least 2 years; in neutral solution they are just stable for several days. In acidic solution nanoparticles are stable for at least several months.

The primary advantage of nanoparticles is their ability to adsorb different drugs because a large variety of compounds can be delivered to the brain. Advantages over other methods of delivering drugs to the brain are that (1) they do not open the BBB, (2) potentially any drug can be delivered (hydrophilic or hydrophobic), and (3) the drug does not need to be altered. Furthermore, nanoparticles are known to be nontoxic. Neither the polymer nor the surfactants have any known toxic effect.

Efficacy of the nanoparticle approach has been demonstrated by tests with an analgesic peptide, dalargin, as well as with several other drugs, including cytostatic agents (e.g., doxorubicin). Dalargin normally does not cross the BBB because it only has an analgesic effect when injected directly into the brain. However, when bound to nanoparticles, a marked analgesia is seen when it is given peripherally.

In summary, nanoparticles are a very useful and universal method to deliver drugs to the brain. Industrial applications of the nanosphere technology would have several benefits:

- Nanoparticles deliver drugs to the brain that normally do not cross the blood–brain barrier.
- They reduce peripheral side effects of (approved) drugs that cross the BBB by increasing the relative dose of drugs reaching the brain;
- Nanoparticles can also be used as a screening tool. Delivering drug candidates to the brain by nanosphere technology for initial screening of CNS activity obviates direct CNS injections. It also reduces the need for altering drugs to allow their passage through the BBB, which decreases drug development costs significantly. Furthermore, drugs that come off patent protection can be protected again when used in combination with new drug delivery tools such as nanoparticles.

Nanoparticles thus open new possibilities for the treatment of disorders of the brain that were previously inconceivable. The value of nanotechnology for medicine is therefore obvious. We are confident that nanotechnology will make a major contribution to the advancement of drug treatment by helping drugs to be targeted more efficiently to specific organs, such as the brain. This may also be a means to attack previously untreatable disorders such as brain tumors and other neurodegenerative diseases.

GLOSSARY

Adsorption Binding of substances to a solid surface.

Aldehyde [alcohol + New Lat. *dehydrogenatum*, dehydrogenated], any of a class of organic compounds that contain the carbonyl group, and in which the carbonyl group is bonded to at least one hydrogen; the general formula for an aldehyde is RCHO, where R is hydrogen or an alkyl or aryl group.

Bioavailability After intraveneous or oral administration a drug is distributed throughout the body in different concentrations in several organs.

Chromatography Resolution of a chemical mixture into its component compounds by passing it through a system that retards each compound to a varying degree; a system capable of accomplishing this is called a chromatograph.

Colloid A mixture in which one substance is divided into minute particles (called colloidal particles) and dispersed throughout a second substance. The mixture is also called a colloidal system, colloidal solution, or colloidal dispersion. Familiar colloids include fog, smoke, homogenized milk, and ruby-colored glass.

Cryoprotector Compound such as a sugar that prevents the nanoparticles from the formation of agglomerates.

Denaturation Term used to describe the loss of native, higher-order structure of protein molecules in solution. Most globular proteins exhibit complicated three-dimensional folding described as secondary, tertiary, and quarternary structures.

Dialysis Transfer of solute (dissolved solids) across a semipermeable membrane. Strictly speaking, dialysis refers

only to the transfer of the solute; transfer of the solvent is called osmosis. Dialysis is frequently used to separate different components of a solution.

Dispersion Mixture in which fine particles of one substance are scattered throughout another substance. A dispersion is classified as a suspension, colloid, or solution. Generally, the particles in a solution are of molecular or ionic size.

Enzyme Biological catalyst; enzymes are proteins that accelerate the rates of reactions while experiencing no permanent chemical modification as a result of their participation.

Evaporation Change of a liquid into vapor at any temperature below its boiling point. For example, water, when placed in a shallow open container exposed to air, gradually disappears, evaporating at a rate that depends on the amount of surface exposed, the humidity of the air, and the temperature.

Free radical A molecule or atom that contains an unpaired electron but is neither positively nor negatively charged. Free radicals are usually highly reactive and unstable. They are produced by homolytic cleavage of a covalent bond.

Ion A neutral atom or group of atoms becomes an ion by gaining or losing one or more electrons or protons. Because the electron and proton have equal but opposite unit charges; the charge of an ion is always expressed as a whole number of unit charges and is either positive or negative.

Isotherm Line drawn on a map connecting points of equal temperature. Each point reflects one temperature reading or an average of several readings over a period of time.

Lipid A broad class of organic products found in living systems. Most are insoluble in water but soluble in nonpolar solvents. The definition excludes the mineral oils and other petroleum products obtained from fossil material. Major classes of lipids include the fatty acids, the glycerol-derived lipids, the sphingosine-derived lipids, the steroids and their derivatives, the terpenes and their derivatives, certain aromatic compounds, and long-chain alcohols and waxes.

Phospholipid Lipid that in its simplest form is composed of glycerol bonded to two fatty acids and a phosphate group. The resulting compound called phosphatidic acid contains a region (the fatty acid component) that is fat-soluble along with a region (the charged phosphate group) that is water-soluble.

Polymerization Reaction to create a chemical compound with high molecular weight consisting of a number of structural units linked together by covalent bonds.

Precipitation A process in which a solid is separated from a suspension, sol, or solution. In a suspension such as sand in water the solid spontaneously precipitates (settles out) on standing. In a sol the particles are precipitated by coagulation.

Surfactant A surface-active compound with a hydrophilic and a lipophilic moiety.

Suspension Mixture of two substances, one of which is finely divided and dispersed in the other. Common suspensions include sand in water, fine soot or dust in air, and droplets of oil in air.

REFERENCES

1. M. W. Bradbury, *Exp. Physiol.* 78, 453 (1993).
2. A. M. Butt, H. C. Jones, and N. J. Abbott, *J. Physiol.* (*London*) 429, 47 (1990).
3. H. C. Jones, R. F. Keep, and A. M. Butt, *Prog. Brain Res.* 91, 123 (1990).
4. F. Joo, *Neurochem. Int.* 23, 499 (1993).
5. W. M. Pardridge, D. Triguero, J. Yan, and P. A. Cancilla, *J. Pharmacol. Exp. Ther.* 253, 844 (1990).
6. M. G. Donelli, M. Zucchetti, and M. D'Incalci, *Cancer Chemother. Pharmacol.* 30, 251 (1992).
7. C. Chesne, M. P. Dehouck, P. Jolliet Riant, F. Bree, J. P. Tillement, B. Dehouck, J. C. Fruchard, and R. Cecchelli, *Adv. Exp. Biol.* 331, 113 (1993).
8. C. Crone and S. P. Olsen, *Brain Res.* 241, 49 (1982).
9. M. P. Dehouck, S. Merèsse, P. Delorme, J. C. Fruchart, and R. Cecchelli, *J. Neurochem.* 54, 1789 (1990).
10. F. E. Arthur, R. R. Shivers, and P. D. Bowman, *Dev. Brain Res.* 36, 155 (1987).
11. J. H. Tao-Cheng, Z. Nagy, and M. W. Brightman, *J. Neurosci.* 7, 3293 (1987).
12. P. A. Stevard and M. L. Wiley, *Dev. Biol.* 84, 183 (1981).
13. P. A. D'Amore, in "Cell Culture Techniques in Heart and Vessel Research" (H. M. Piper, Ed.), p. 299. Springer-Verlag, Berlin, 1990.
14. L. E. Stramm, W. Li, G. D. Aguirre, and J. H. Rockey, *Exp. Eye Res.* 44, 98 (1987).
15. D. Shepro and M. L. Morel, *FASEB J.* 7, 1031 (1993).
16. R. N. Frank, S. Dutta, and M. A. Marcini, *Invest. Ophthal. Visual Sci.* 28, 1086 (1980).
17. V. Nehls and D. Drenckhahn, *Histochemistry* 99, 1 (1993).
18. A. Frey, B. Meckelein, H. Weiler-Güttler, B. Möckel, R. Flach, and H. G. Gassen, *Eur. J. Biochem.* 202, 421 (1991).
19. W. Risau, A. Dingler, U. Albrecht, M. P. Dehouck, and R. Cecchelli, *J. Neurochem.* 58, 667 (1992).
20. H. K. Kimelberg and M. D. Norenberg, *Spektr. Wiss.* 6, 52 (1989).
21. G. W. Goldstein, *Ann. N.Y. Acad. Sci.* 529, 31 (1988).
22. W. H. Oldendorf, *Am. J. Physiol.* 224, 1450 (1973).
23. W. H. Oldendorf, *Eur. Neurol.* 6, 49 (1972).
24. T. Terasaki, S. Takakuwa, S. Moritani, and A. Tsuji, *J. Pharmacol. Exp. Ther.* 258, 932 (1991).
25. T. Terasaki, Y. S. Kang, T. Onishi, and A. Tsuji, *J. Pharm. Pharmacol.* 43, 172 (1991).
26. H. Takanaga, I. Tamai, S. Inabe, Y. Sai, H. Higashiga, H. Yamamoto, and A. C. Tsuji, *Biochem. Biophys. Res. Commun.* 217, 370 (1995).
27. A. Kurihara, H. Suzuki, Y. Sawada, Y. Sugiyama, T. Iga, and M. Hanano, *J. Pharmacol. Sci.* 76, 759 (1987).
28. W. M. Pardridge, R. Sakiyama, and G. Fierer, *J. Clin. Invest.* 71, 900 (1982).
29. Y. S. Kang, T. Terasaki, T. Onishi, and A. Tsuji, *J. Pharmacobiol. Dyn.* 13, 353 (1990).
30. M. Shimon, Y. Egozi, Y. Kloong, M. Sokolovski, and S. Cohen, *J. Neurochem.* 50, 1719 (1988).
31. E. M. Conford, L. D. Braun, and W. H. Oldendorf, *J. Neurochem.* 30, 299 (1978).
32. M. Yamazaki, H. Fukuoka, O. Nagata, H. Kato, T. Terasaki, and A. Tsuji, *Biol. Pharm. Bull.* 17, 676 (1994).
33. W. A. Banks, A. J. Kastin, and C. M. Barrera, *Pharm. Res.* 8, 1345 (1991).
34. R. Kannan, J. F. Kuhlenkamp, E. Jeandidier, H. Trinh, M. Ookhtens, and N. Kaplowitz, *J. Clin. Invest.* 85, 2009 (1990).
35. R. Kannan, J. F. Kuhlenkamp, M. Ookhtens, and N. Kaplowitz, *J. Pharmacol. Exp. Ther.* 263, 964 (1992).
36. A. K. Kumagi, J. Eisenberg, and W. M. Pardrige, *J. Biol. Chem.* 262, 15214 (1987).

37. W. M. Pardridge, D. Triguero, and J. Buciak, *J. Pharmacol. Exp. Ther.* 251, 821 (1989).
38. W. M. Pardridge, "Peptide Drug Delivery to the Brain." Raven Press, New York (19xx).
39. K. R. Smith and R. T. Borchardt, *Pharm. Res.* 6, 466 (1989).
40. W. M. Pardridge, J. L. Buciak, and P. M. Frieden, *J. Pharmacol. Exp. Ther.* 259, 66 (1991).
41. W. M. Pardridge, A. K. Kumagai, and J. B. Eisenberg, *Biochem. Biophys. Res. Commun.* 146, 307 (1987).
42. V. A. Levin, *J. Med. Chem.* 23, 682 (1980).
43. W. M. Pardridge, D. Tiguero, J. Yang, and P. A. Cancilla, *J. Pharmacol. Exp. Ther.* 253, 884 (1990).
44. J. L. Biedler and H. Riehm, *Cancer Res.* 30, 1174 (1970).
45. R. L. Juliano and V. Ling, *Biochim. Biophys. Acta* 455, 152 (1989).
46. G. Bradley, P. F. Juranka, and V. Ling, *Biochim. Biophys. Acta* 948, 87 (1989).
47. T. Skovsgaard, *Cancer Res.* 38, 4722 (1978).
48. P. Gross, Y. B. Neriah, J. M. Croop, and D. E. Housman, *Nature* (*London*) 323, 728 (1986).
49. A. R. Safa, *Proc. Natl. Acad. Sci. U.S.A.* 85, 7187 (1989).
50. I. Tamai and A. R. Safa, *J. Biol. Chem.* 266, 16796 (1991).
51. A. R. Safa, C. J. Glover, M. B. Meyers, J. L. Biedler, and R. L. Felsted, *J. Biol. Chem.* 261, 6137 (1986).
52. B. A. Sabel, A. Freese, and M. J. During, *Adv. Neurol.* 53, 521 (1990).
53. W. H. Oldendorf, *Am. J. Physiol.* 221, 1629 (1971).
54. J. B. M. M. Van Bree, A. G. De Boer, M. Danhof, and D. D. Breimer, *Pharm. World Sci.* 15, 2 (1993).
55. M. Hiesmayr, H. Dirnberger, A. Aloy, K. Heimberger, A. Horaczek, and B. Branstatter, *Schweiz. Med. Wochenschr.* 117, 450 (1987).
56. E. A. Neuwelt, M. Glasberg, E. Frenkel, and P. Barnett, *Ann. Neurol.* 14, 316 (1983).
57. W. M. Pardridge, *Endocr. Rev.* 7, 314 (1986).
58. W. A. Jeffries, M. R. Brandon, S. V. Hunt, A. F. Williams, K. C. Gatter, and D. Y. Mason, *Nature* (*London*), 312, 162 (1984).
59. J. Kreuter, R. N. Alyautin, D. A. Kharkevich, and A. A. Ivanov, *Brain Res.* 674, 171 (1995).
60. J. Kreuter, V. E. Petrov, D. A. Kharkevich, and R. N. Alyautdin, *J. Controlled Release* 49, 81 (1997).
61. U. Schroeder and B. A. Sabel, *Brain. Res.* 710, 121 (1996).
62. U. Schroeder, P. Sommerfeld, and B. A. Sabel, *Peptides* 19, 777 (1998).
63. U. Schroeder, S. Ulrich, and B. A. Sabel, *J. Pharm. Sci.* 87, 1305 (1998).
64. U. Schroeder, B. A. Sabel, and H. Schroeder, *Life Sci.* 66, 495 (2000).
65. J. Kreuter, *J. Anat.* 189, 503 (1996).
66. J. Kreuter, "Encyclopedia of Pharmacy Technology," 165 (1994).
67. J. Kreuter, in "Specialized Drug Delivery Systems" (P. Tyle, Ed.), p. 257. Marcel Dekker, New York, 1990.
68. G. Birrenbach and P. P. Speiser, *J. Pharm. Sci.* 65, 1763 (1976).
69. N. Al Khouri Fallouh, L. Roblot-Treubel, H. Fessi, J. P. Devissaguet, and F. Puisieux, *Int. J. Pharm.* 28, 125 (1986).
70. T. R. Tice and R. M. Gilley, *J. Controlled Release* 2, 343 (1985).
71. H.-J. Krause, A. Schwartz, and P. Rohdewald, *Int. J. Pharm.* 27, 145 (1985).
72. W. P. Yu, J. P. Wong, and T. M. S. Chang, *J. Microencaps.* 15, 515 (1998).
73. H. Fessi, F. Puisieux, J. Devissaguet, N. Ammoury, and S. Benita, *Int. J. Pharm.* 55, R1 (1989).
74. N. Ammoury, H. Fessi, J. P. Devissaguet, F. Puisieux, and S. Benita, *J. Pharm. Sci.* 79, 763 (1990).
75. I. Zolle, F. Hosain, B. A. Rhodes, and H. N. Wagner, Jr., *J. Nucl. Med.* 11, 379 (1970).
76. U. Scheffel, B. A. Rhodes, T. K. Natarajan, and H. N. Wagner, Jr., *J. Nucl. Med.* 13, 498 (1972).
77. J. J. Burger, E. Tomlinson, and J. W. Mulder, *Int. J. Pharm.* 23, 333 (1985).
78. P. Sommerfeld, U. Schroeder, and B. A. Sabel, *Int. J. Pharm.* 155, 201 (1997).
79. P. Sommerfeld, B. A. Sabel, and U. Schroeder, *J. Microencaps.* 17, 69 (2002).
80. P. Sommerfeld, U. Schroeder, and B. A. Sabel, *Int. J. Pharm.* 164, 113 (1998).
81. S. J. Douglas, L. Illum, and S. S. Davis, *J. Colloid. Interface Sci.* 103, 154 (1985).
82. U. E. Berg, J. Kreuter, and P. P. Speiser, *Pharm. Ind.* 48, 75 (1986).
83. H. S. Yalabik-Kas, J. Kreuter, A. A. Hincal, and P. P. Speiser, *J. Microencaps.* 3, 71 (1986).
84. J. Kreuter, in "Topics in Pharmaceutical Sciences" (D. D. Bremer and P. P. Speiser, Eds.), p. 359. Elsevier, Amsterdam, 1983.
85. T. Harmia, P. P. Speiser, and J. Kreuter, *J. Microencaps.* 3, 3 (1986).
86. U. E. Berg, J. Kreuter, P. P. Speiser, and M. Solvia, *Pharm. Ind.* 48, 75 (1986).
87. H. Kopf, R. K. Joshi, M. Solvia, and P. P. Speiser, *Pharm. Ind.* 39, 993 (1977).
88. L. Illum, M. A. Khan, E. Mak, and S. S. Davis, *Int. J. Pharm.* 30, 17 (1986).
89. A. El-Egakey and P. P. Speiser, *Pharm. Acta Helv.* 57, 236 (1982).
90. P. Beck, D. Scherer, and J. Kreuter, *J. Microencaps.* 7, 491 (1990).
91. P. K. Gupta, C. T. Hung, and D. G. Perrier, *Int. J. Pharm.* 33, 137 (1986).
92. R. Diepold, J. Kreuter, J. Himber, R. Gurny, V. H. Lee, J. R. Robinson, M. F. Saettone, and O. E. Schnaudigel, *Graefes Arch. Clin. Exp. Ophthalmol.* 227, 188 (1989).
93. P. Couvreur, B. Kante, M. Roland, and P. J. Speiser, *Pharm. Sci.* 68, 1521 (1979).
94. P. R Lockman, J. R. Mumper, M. A. Khan, and D. D. Allen, *Drug Dev. Ind. Pharm.* 28, 1 (2002).
95. J. Kreuter, C. G. Wilson, J. R. Fry, P. Paterson, and J. H. Ratcliffe, *J. Microencaps.* 1, 253 (1984).
96. B. Kante, P. Couvreur, G. Dubois-Krack, C. D. E. Meester, P. Guiot, M. Roland, M. Mercier, and P. Speiser, *J. Pharm. Sci.* 71, 786 (1982).
97. L. Grislain, P. Couvreur, V. Leanaerts, M. Roland, D. Deprez-Decampeneere, and P. Speiser, *Int. J. Pharm.* 15, 335 (1983).
98. J. Kreuter, *Pharm. Acta Helv.* 58, 217 (1983).
99. J. Kreuter, *Adv. Drug Del. Rev.* 47, 65 (2001).
100. P. Couvreur, L. Grislain, V. Lenaerts, P. Brasseur, P. Guiot, and A. Biernacki, in "Polymeric Nanoparticles and Microspheres" (P. Guoit and P. Couvreur, Eds.), p. 27. CRC Press, Boca Raton, FL, 1986.
101. F. Leonard, R. K. Kulkarni, G. Brandes, J. Nelson, and J. J. Cameron, *J. Appl. Polymer. Sci.* 10, 259 (1966).
102. R. Alyautdin, D. Gothier, V. Petrov, D. A. Kharkevich, and J. Kreuter, *Eur. J. Pharm. Biopharm.* 41, 44 (1995).
103. P. Ramge, J. Kreuter, and B. Lemmer, *Chronobiol. Int.* 17, 767 (1999).
104. R. N. Alyautdin, V. E. Petrov, K. Langer, A. Berthold, D. A. Kharkevich, and J. Kreuter, *Pharm. Res.* 14, 325 (1997).
105. R. N. Alyautdin, E. B. Tezikov, P. Ramge, D. A. Kharkevich, D. J. Begley, and J. Kreuter, *J. Microencaps.* 15, 67 (1998).
106. J. Darius, F. P. Meyer, B. A. Sabel, and U. Schroeder, *J. Pharm. Pharmacol.* 52, 1043 (2000).
107. J. Darius, *J. Chromatogr., B* 682, 67 (1996).
108. A. E. Gulyaev, S. E. Gelperina, I. N. Skidan, A. S. Antropov, G. Y. Kivman, and J. Kreuter, *Pharm. Res.* 16, 1564 (1999).
109. P. Couvreur, B. Kante, L. Grislain, M. Roland, and P. Speiser, *J. Pharm. Sci.* 71, 790 (1982).
110. L. Illum and S. S. Davis, *J. Pharm. Sci.* 72, 1086 (1983).
111. L. Illum, S. S. Davis, R. H. Müller, E. Mak, and P. West, *Life Sci.* 40, 367 (1987).

112. S. D. Tröster, U. Müller, and J. Kreuter, *Int. J. Pharm.* 61, 85 (1990).
113. S. E. Gelperina, Z. S. Smirnova, A. S. Khalanskiy, I. N. Skidan, I. A. Bobruskin, and J. Kreuter, "Proceedings of the 3rd World Meeting APV/APGI," Berlin, 2000, p. 441.
114. S. K. Pulfer and J. M. Gallo, *J. Drug Target.* 6, 215 (1998).
115. M. Lück, Ph.D. Thesis, Freie Universität, Berlin, 1997, pp. 14, 137.
116. J. Kreuter, D. Shamenkov, V. Petrov, P. Ramge, K. Cychutek, C. Koch-Brandt, and R. Alyautdin, *J. Drug Target.* 10, 317 (2002).
117. C. Wilson, M. R. Wardell, K. H. Weisgraber, R. W. Mahley, and D. A. Agard, *Science* (*Washington, DC*) 252, 1817 (1991).
118. C.-Y. Yang, S.-H. Cheng, S. H. Gianturco, W. H. Bradley, J. T. Sparrow, H. DeLooft, M. Rosseneu, Z.-S. Lee, Z.-W. Gu, A. M. Gotto, and L. Chan, *Nature* (*London*) 323, 738 (1986).
119. B. Dehouck, L. Fenart, M.-P. Dehouck, A. Pierce, G. Torpier, and R. Cecchelli, *J. Cell Biol.* 138, 877 (1997).
120. B. Dehouck, M. P. Dehouck, J. C. Fruchart, and R. Cecchelli, *J. Cell Biol.* 126, 465 (1994).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanoparticle Layers in Multilayers

Diana Nesheva

Bulgarian Academy of Sciences, Sofia, Bulgaria

CONTENTS

1. INTRODUCTION

Over the past several decades, there has been rapidly increasing interest in the preparation of nanometer-sized materials such as superlattices, quantum wires, and quantum dots. This enhanced interest is justified because these materials show a great variety of new properties with respect to three-dimensional ones. Nanostructured or nanophase materials can be classified into four categories according to the shape of their structural constituents: nanophase powders, nanostructured films (including single-layer, multilayer, composite film, compositionally graded film, etc.), monolitic nanostructured materials, and nanostructured composite. The nanostructured materials can contain crystalline, quasicrystalline, and/or amorphous phases.

The search for ultra-thin materials (nanostructured films), in particular, semiconductors, can be traced quite far back (for a review of early work up to 1975, see [1–3]). However, the motivation for their production went up sharply when new types of devices were predicted [4, 5]. Studies of ultra-thin semiconductor layers and multilayer structures containing ultra-thin films have since then increased progressively. Single and multiple quantum wells, single- and double-barrier tunneling structures, incoherent multilayer tunneling structures, etc. have been prepared [6–10] as some of them have been used in production of various devices such as mobile phones, semiconductor photomultipliers, and lasers, etc. [11, 12]. It is important to notice that in crystalline superlattices (periodic multilayer structures built of two consecutively deposited materials) disorder and scattering must be low enough to allow building up of coherent superlattice bandstates and, thus, to prevent destruction of the phase coherence by interface fluctuations. This requirement means that for building up crystalline superlattices, pairs of materials have to be used with a difference in the lattice constants smaller that 1%. Thus, only a limited number of pairs can be applied in the superlattice preparation.

In the beginning of the 1980's, it was shown [13] that superlattices could also be fabricated from amorphous semiconductors (a-Si:H, a-Ge:H, SiO_x, SiN_x, etc.). The lack of long-range order in the structure of amorphous materials induces a great number of defects such as bond-length and bond-angle fluctuations, dangling and floating bonds, etc. This fact, along with the flexibility of amorphous materials, strongly reduces the requirements for matching lattice constants of the constituent materials and makes possible application of a great variety of materials in the fabrication of amorphous multilayers (a-MLs). Regular a-MLs, based on a-Si:H and a-Ge:H, have been prepared by a variety of rather simple and cheaper set-ups (glow-discharge [13], magnetron sputtering [14], plasma- and photochemically enhanced chemical vapor deposition [15–17] techniques), as compared to molecular beam epitaxy [18, 19] applied for crystalline MLs. It has been shown that applying bandgap engineering, new properties can be achieved, based both on classical effects, connected mainly with the influence of the interlayer barriers [20–25], and quantum-size effect (QSE) when the layers are sufficiently thin and the period (or the well layer width) is comparable with the de Broglie wavelengths of the charge carriers (electrons and holes) [20, 26–30]. Four-fold coordinated a-Si:H, a-Ge:H, and similar materials keep the rigid tetrahedrally bonded structure of the respective elemental crystalline semiconductors. In chalcogenide materials coordination number (that is, the number of the nearest neighbors) varies between 2 (for a-S, a-Se, a-Te, and their alloys) and <3 for $Ge_xS(Se)_{100-x}$ compounds. This low-coordination number makes amorphous chalcogenides more flexible than a-Si:H, a-Ge:H, and similar four-fold coordinated materials. Thus, in chalcogenide a-MLs, the level of interface-induced network distortion (and, hence, the concentration of interface defects) is lower than in a-Si, a-Ge-based MLs [23].

The main peculiarity of nanoparticles in nanocomposite films and bulk nanostructured materials is the presence of a large fraction of atoms located at the surface of nanoparticles embedded in a matrix. The large fraction of

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (105–123)

low-coordinated atoms at the surfaces of semi-isolated nanocrystals or within the grain boundaries of nanometer-sized composite or polycrystalline materials determines many of their physical, chemical, and mechanical properties [32–35]. The electronic eigenstates and phonons are described by periodic wave functions localized within the nanocrystal. The quantum confinement results in an increase of the bandgap and formation of discrete states at the band edges of semiconductor nanocrystals. The selection rules for optical transitions are relaxed owing to the scattering of the electronic eigenstates and phonons at the grain boundaries, which leads to their mixing and increase of transition probabilities for light absorption and emission. The localization of phonons in the nanocrystals is seen as a shift of the bands in their Raman scattering spectra.

In this chapter, the focus is on fabrication of nanocrystalline materials by nonepitaxial vapor deposition techniques and post-treatment of regular (amorphous or nanocrystalline/amorphous) multilayer structures. Using this approach one can produce either continuous or discontinuous nanocrystalline layers from a great variety of materials. The main advantage of this approach is the possibility for a precise control of the layer thickness during the ML deposition procedure. This makes it possible for the fabrication of nanocrystals (NCs) with desired sizes. In Section 2, a brief description of vapor deposition techniques applied for fabrication of amorphous and crystalline MLs is given. The modes of film growth are also considered. In Section 3, attention is paid to preparation of high-quality periodic amorphous multilayer, thermally and laser-beam induced crystallization of their constituent layers. Special emphasis is put on changes in the crystallization temperature of two-dimensional amorphous materials. Crystallization-induced changes in MLs and material intermixing upon annealing is also considered in this section. In Section 4, preparation of discontinuous metal and semiconductor nanoparticle films is described by means of the multilayer approach and possible mechanisms of nanoparticle formation are discussed.

2. VAPOR DEPOSITION OF THIN FILMS

Vapor processing methods seem to be the only ones that can provide highly pure materials with structural control at the atomic level or nanometer scale [36]. Moreover, vapor processing routs can produce ultra-fine powders, multilayers, and functionally graded materials, as well as composite materials with well-controlled dimensions and unique structures at a lower processing temperature. Vapor processing techniques used for fabrication of nanophase materials can be classified into physical vapor deposition (PVD), chemical vapor deposition (CVD), and aerosol-based processes and flame-assisted deposition methods [36, 37]. The first two have been widely applied for preparation of nanostructured films. PVD can be subdivided into evaporation, sputtering, and ion-plating processes based on the different ways of generating the gaseous species (Fig. 1). The classification of CVD processes based on the different heating methods can be as follows: (1) thermally activated, (2) photo-assisted, (3) plasma-assisted, (4) metalorganic, and (5) atomic layer epitaxy (Fig. 1).

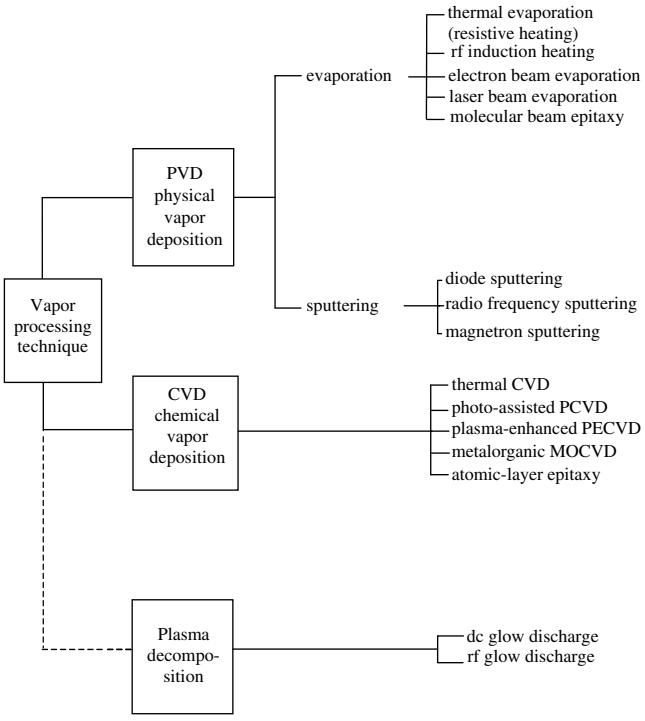


Figure 1. Variants of vapor deposition techniques for the fabrication of nanostructured films.

2.1. PVD Processes

In PVD processes, the particles to be deposited are transformed into the gaseous state by a physical process, either thermal evaporation or an impact process. By heating the kinetic energy of atoms or molecules of a solid or liquid, increases and at high enough temperatures, a significant number of atoms or molecules overcomes the separation energy and evaporate. There are various possibilities for evaporating the material to be deposited. The heat for evaporation can be delivered by resistive heating of a refractory metal, as well as by a laser or electron beam, high-frequency induction, or arc-evaporation. Generally, PVD methods can be mainly distinguished by the size of the ejected particles, by the degree of ionization, and by energy distribution. Moreover, the properties of growing films can also be influenced by the substrate temperature, the gas pressure, and an extra bombardment with high-energy neutral atoms or ions of an inert gas.

2.1.1. Evaporation

The most convenient thermal evaporation method uses resistive heating of a crucible. It works well if no chemical reaction or substantial alloying of the crucible with the evaporant takes place, and the evaporating material has a significantly lower melting point than the crucible. The heating of the evaporant by high-frequency induction is also applicable but only for materials with certain electrical conductivity. In this technique, the evaporating material levitates above the crucible and no pollution from crucible is introduced.

On application of an electron beam gun, the electrons emitted from a tungsten cathode are accelerated by a high

voltage (5–10 kV) and are focused on the substance to be evaporated. The kinetic energy of the electrons is transformed into thermal energy so that the material in the target melts and evaporates in vacuum. There are two advantages in using an electron gun compared with the resistive heating of a crucible: the evaporation rate is significantly higher and there is no alloying of the evaporant with the cold crucible.

The laser ablation technique uses a pulse laser beam of high power for ejecting particles into vacuum. The ejection process is very complicated: a large portion of the ejected particles are ionized and the evaporating beam consists of components of different velocities. Various ejection processes of the target material have been discussed [37]:

1. striking out of flocks by thermal stress (exfoliational sputtering);
2. ejection of fine droplets by a volume change that arises from thermal expansion of the liquid with melting (hydrodynamic sputtering);
3. production of an electronic avalanche similar to that in electric breakthrough and ionizing collisions of electrons (in this case, a highly excited dense plasma in front of the target surface exists);
4. transmission of momentum from the beam on the target surface (collisional sputtering);
5. normal thermal evaporation.

An advantage of laser evaporation is that because of its high energy, alloys can be deposited without significant change of composition as in flash evaporation technique. This process can also be applied reactively.

2.1.2. Sputtering

If ions or atoms with energy several kiloelectronvolts bombard a solid surface, single atoms, molecules, or clusters are produced which are deposited on the surrounding wall of the vacuum chamber [36, 37]. The ions of a noble gas are mostly applied to bombard the solid surface. In standard sputtering equipment, the well-known glow discharge is used as the ion source. It forms if one puts an electric field of several hundreds or a thousand volts between two electrodes which stand separately in a vessel evacuated to about 10^{-1} mbar. In the glow discharge, not only neutral atoms but also ions and electron exist. The electrons are accelerated against the anode by the applied electric field and produce additional ions on the way of impact ionization. The ions are accelerated against the cathode and produce these secondary electrons which produce again so many new ions in the gas by collisional ionization that the ion loss on the cathode is compensated. This means that the discharge is self-sustaining. By the impact of positive gas ions on the cathode, the atoms of the cathode material are ejected and condense on the substrates forming the film. The particles sputtered from the cathode are more than 95% neutral and their average energy is considerably higher (5–20 eV) than that of thermally evaporated atoms (<1 eV). When applying the sputtering technique, one should bear in mind that some further processes take place:

1. the growing film is bombarded by neutral atoms of the inert gas;
2. if the substrate is not properly grounded, it will be bombarded either by ions of the inert gas or secondary electrons that yields heating of the substrate;
3. in case of reactive sputtering, in addition to the mentioned two processes, the reaction rate between the depositing atom and reactive gas on the substrate will be further increased.

Sputtering techniques include diode sputtering, radio frequency (rf) sputtering, and magnetron sputtering. Direct-current (dc) glow discharge sputtering can only be used for conductive targets, while any target regardless of its conductivity can be sputtered using rf sputtering. Magnetron or triode sputtering is used to provide a higher current in the relatively inefficient ion source than that provided by glow discharge. In this case, a higher ionization of the plasma is obtained by introducing electrons into the plasma from a thermionic emitter [38].

Another method for the deposition of films, which is on the border between reactive sputtering and CVD methods, is the decomposition of gases in plasma. With this method, no evaporation source and target are used. The material to be decomposed is admitted as a component of a gaseous compound to the reaction chamber. In the plasma, the molecules of the compound decompose and the desired component is deposited on the substrate. Frequently, a negative bias voltage is put on the substrate so that the growing film is bombarded with ions. An rf or dc voltage may be applied for sustaining the glow discharge. This technique is widely used for deposition of amorphous silicon (a-Si:H) by plasma decomposition of SiH_4.

2.2. Chemical Vapor Deposition

CVD involves thermal dissociation and/or chemical reaction (thermal decomposition (pyrolysis) oxidation, reduction, hydrolysis, nitridation, etc.) of the gaseous reactants on or near the vicinity of a heated surface to form stable solid products [37]. The deposition involves homogeneous and/or heterogeneous chemical reactions leading to the formation of films or powders. During the CVD film deposition, the active gaseous species is generated and transported into the reactor chamber where the gaseous precursor undergoes gas phase reactions forming an intermediate phase. The species of this phase are then absorbed onto the heated substrate, where heterogeneous reactions take place producing the deposit and byproducts. The latter are diffused away from the substrates and transported out from the deposition chamber. The deposit is diffused along the heated substrate surface forming embryos for subsequent growth of the film.

Photo-assisted chemical vapor deposition (PCVD) is a variant of the CVD process that relies on absorption of laser radiation in the substrate to raise its temperature and cause thermal decomposition/chemical reaction of the gaseous precursor to form the required products. Plasma-enhanced CVD (PECVD) uses electron energy (plasma) as the activation method. By supplying electrical power at sufficiently high voltage to a gas at reduced pressure (<10 mbar), breakdown of the gas occurs and generates a glow discharge plasma consisting of electrons, ions, and electronically excited species. The plasma ionizes and decomposes

the reactant gases at a low temperature. Chemical vapor deposition can also be classified according to the type of precursor used. The use of a metalorganic precursor, for example, has led to the development of metalorganic chemical vapor deposition (MOCVD). Atomic layer epitaxy (ALE) is a variant of the CVD process that involves surface deposition for the controlled growth of epitaxial films and fabrication of tailored molecular structures on solid surfaces [39]. "Monoatomic" layers can be grown in sequence by sequentially saturating surface reactions.

2.3. Modes of Film Growth

To understand the deposition of a film on a substrate, it is necessary to consider the nucleation and the later growth separately. Depending principally on the material parameters of the substrate and the deposited material, as well as the supersaturation of the vapor, three different modes of film growth have been observed:

1. If the interaction energy W_{SD} (adhesion) between atoms of the substrate and deposit is much larger than the interaction energy W_{DD} (cohesion) between atoms of the deposit ($W_{SD} \gg W_{DD}$), a subsequent layer will grow after the previous one is completed (layer growth (LG)).
2. If the cohesion energy is larger than the adhesion ($W_{DD} > W_{SD}$), island growth (IG) takes place.
3. The Stranski-Krastanow (SK) growth is a mixture of LG and IG [40].

Stranski-Krastanow growth occurs if the interaction energies have very different dependencies on the distance between the atoms. The increase in the supersaturation leads to a more two-dimensional growth. A drastically increased supersaturation can change the growth mode from IG to LG.

The following processes take place on the substrate surface during the initial IG of films:

(a) the formation of adatoms on the substrate surface by the arrival of the vapor particles;
(b) the diffusion of adatoms on the substrate surface;
(c) the partial desorption of adatoms from the surface into the vacuum;
(d) the statistical nucleation on a substrate surface which is free of defects;
(e) the nucleation on defects;
(f) the growth of nuclei by the capture of adatoms.

For evaporation on an undisturbed surface, the island density increases first by nucleation, goes through a maximum at t_{max}, and then decreases monotonically up to 100% coverage of the substrate surface because of the coalescence of islands. With rising substrate temperature, the nucleation rate and island density decreases and t_{max} increases. A reduction in the arrival rate yields qualitatively a similar result. If the kinetic energy of the vapor atoms that arrive is increased, the nucleation rate and island density increase while t_{max} decreases. With increasing substrate temperatures the mobility of adatoms and clusters increases which favors grain growth, recrystallization, and an increasing order of the microstructure.

Epitaxy means that the growing layer derives its crystalline orientation from the underlying substrate [41]. Generally, the deposition of epitaxial films on single-crystal substrates with evaporation is possible only with a high substrate temperature T_{epi}. It has been observed that for films prepared by sputtering, T_{epi} is less than for films prepared with evaporation of materials. This is due to the higher energy of the particles in sputtering.

Molecular beam epitaxy (MBE) is a refined form of vacuum evaporation. This method involves [18, 42–44] neutral thermal energy molecular or atomic beams impinging on a hot crystalline substrate maintained in an ultra-high vacuum that provides the required source of growth. In the MBE, the substrate temperature and beam intensities of each source are controlled separately as the growth temperature should provide sufficient surface diffusion to allow layer-by-layer growth and well interfaces between layers with minimal bulk interdiffusion. The major drawback of MBE is that it involves the use of expensive and sophisticated equipment with a high running cost. Moreover, production time could be lost because of the long period of time required for the system to be pumped down to the required ultra-high vacuum.

3. CONTINUOUS NANOPARTICLE LAYERS

For device application, precise engineering of nanoparticle size with focus on control in grain boundary and defect density is essential. Fabrication of periodic nanocrystalline/amorphous multilayers is a promising way for production of nanoparticles with desired well-defined sizes. Nanocrystalline/amorphous multilayers have been produced using two main approaches: post-treatment (furnace or rapid thermal annealing, laser beam crystallization) of regular amorphous multilayers or layer-by-layer deposition of nanocrystalline/amorphous layers of similar thickness.

3.1. Preparation of Regular Amorphous Multilayers

Molecular beam epitaxy, as a mature technology, was employed for engineering amorphous/amorphous MLs to atomic precision at the thickness required. Si/SiO_2 periodic MLs were grown [45–47] using the following deposition procedure. A thin Si layer is deposited on substrate and then it is taken out of the vacuum chamber and submitted to 600 s exposure to ultraviolet ozone. This procedure is repeated to create the necessary number of periods. Later, a-Si/SiO_2 periodic MLs have been prepared with electron beam evaporation [48–53] or rf sputtering [54, 55] and *in-situ* plasma oxidation. Plasma-activated gas cell has been used as the oxygen source [48–50]. Plasma-enhanced chemical vapor deposition of a-Si and layer-by-layer plasma oxidation have been used to fabricate a-Si:H/SiO_2 MLs [56]. The thickness of a-Si:H layers has been controlled by the duration time of plasma oxidation of the silicon layers initially deposited. Nanocrystalline Si layers have been grown on SiO_2 by rapid thermal chemical vapor deposition [57].

Regular a-MLs have also been prepared by a consecutive deposition of two materials applying a variety of rather simple and cheaper set-ups, as compared to the MBE normally used for crystalline MLs. Glow-discharge [13, 58, 59], electron beam evaporation [60], plasma- and photochemically enhanced chemical vapor deposition [15–17, 61–65] techniques, and reactive and magnetron sputtering [66–71] have been applied for preparation of a-Si:H- and a-Ge:H-based a-MLs. Physical vapor deposition [23, 26–28, 30, 72–81], laser ablation [20, 82–85], and electron beam evaporation [86] techniques have been used for fabrication of chalcogenide a-MLs. When one applies a nonepitaxial technique for the preparation of a multilayer structure, each layer of this structure (having a thickness between 2 and 20 nm) can be prepared either by one-step (continuous) material deposition or applying a step-by-step approach. Experiments on preparation of a-MLs based on a-Si:H have shown [87] that multilayers prepared by applying the step-by-step deposition of each layer resulted in smoother interfaces and a better periodicity than in the case of continuous deposition of the layers. This approach has also been applied for preparation of selenium-based a-MLs and, thus, high-quality chalcogenide MLs have been produced with smooth and parallel layers [77, 78, 88].

There are few works on layer-by-layer production of nanocrystalline/amorphous MLs. Low-pressure chemical vapor deposition has been used [89–91] to deposit a nanocrystalline nc-Si layer. By controlled, high-temperature thermal oxidation, part of this layer was oxidized to form SiO_2, thus forming a Si/SiO_2 bilayer. This procedure was repeated to build a multilayer structure. Regular multilayer structures of nc-CdSe/SiO_x (Fig. 2), nc-CdSe/GeS_2, nc-Se/SiO_x, and nc-CdSe/ZnSe have also been produced by consecutive thermal evaporation of each material of the respective pair [92–98].

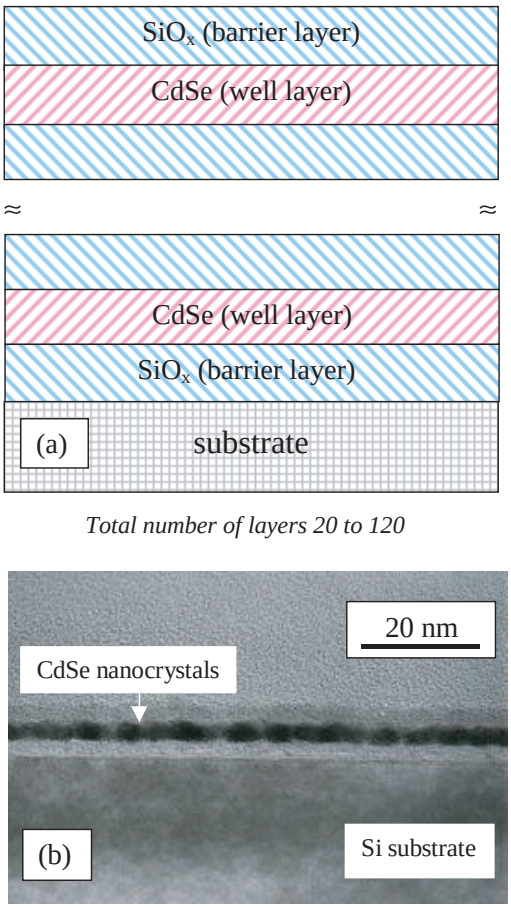


Figure 2. (a) Schematic presentation of a regular multilayer structure: carrier and phonon confinement takes place in the well layers which are made of the material with the narrower optical bandgap. (b) Cross-section view of a SiO_x(5 nm)/CdSe(5 nm)/SiO_x(5 nm) three-layer structure.

3.2. Post-Treatment of Amorphous Multilayers

3.2.1. Thermal Crystallization

Main Features The most convenient way for production of nanocrystalline layers in amorphous MLs is annealing for various times at temperatures around and above the crystallization temperature of the constituent material with a lower crystallization temperature. A standard long-time (>30 min) furnace annealing or rapid thermal annealing (30–60 sec) in an inert atmosphere (Ar, N_2) have been applied to crystallize the layers in a-ML having lower crystallization temperature. Before discussing crystallization, it is interesting to note that the annealing at temperatures close to but below the crystallization temperature (e.g., ~60 min at 350 K for a-Se in a-Se/a-CdSe MLs [88, 99, 100], 20 min at 720 K for a-Ge in a-Ge:H/a-GeN_x MLs [101], etc.) causes a disorder reduction in a-MLs. Besides, upon annealing at temperatures below the crystallization temperature, some authors have observed a thermally induced reduction in the effective interface width, which has been related to negative sublayer interdiffusion [88, 102]. At temperatures higher than the crystallization temperature of the "low-temperature" layers of a-MLs, microcrystal formation takes place (e.g., in a-Ge sublayers of a-Ge:H/a-GeN_x [101], a-Ge:H/a-SiN_x:H [103], a-Si:H/a-Ge:H [104] MLs and a-Se in a-Se/a-CdSe [100, 105], and a-Se/a-As_2S_3 MLs [106]).

The crystallization process in progressively annealed a-MLs has been investigated using differential scanning colorimetry, small- and high-angle X-ray diffraction (SAXRD and HAXRD), as well as high-resolution transmission electron microscopy (HRTEM) and Raman scattering techniques. Earliest investigations of thermal properties of a-MLs (Se/As_2Se_3) [72] performed using differential scanning colorimetry showed a different annealing behavior in pure Se thin films having a thickness of 2 μm and Se/As_2Se_3 MLs. In the temperature range 270 K–470 K, an exothermal peak of crystallization and two endothermal peaks of glass transition and melting were observed in Se thin films. In Se/As_2Se_3 MLs with Se layer thickness <30 nm, no crystallization peak was observed in this temperature range. This was the first indication for some increase in the crystallization temperature of two-dimensional amorphous films. Later, detailed investigations were performed on chalcogenide [88, 95, 99, 100, 105–110], a-Si [102, 111–113], and a-Ge [66, 101, 103, 104, 114, 115] based a-MLs which proved the increase of crystallization temperature T_c. This T_c increase will be discussed in detail later. Furthermore,

upon rapid thermal annealing of a-Si/SiO_2 MLs, in which the a-Si layer thickness ranged from 4 nm to 25 nm, crystallization has been observed at temperatures 1070 K to 1120 K [63], while in a-Si single films, crystallization temperature is near 970 K [116]. Even at such high temperatures, a small amount of residual amorphous tissue (5–10% of the volume) remained in a-Si/SiO_2 MLs after this rapid thermal annealing. To achieve complete crystallization, an additional 30 min furnace annealing at 1320 K has been used [63], as the temperature was raised from 1070 K to 1320 K using a slow ramp of 10 K/min.

Nanocrystallite Shape and Size First studies of the shape of the Si NCs in nc-Si/SiO_2 MLs have shown [63] that nanocrystal are almost spherical with their packing density controlled by the thickness of SiO_2. Later HRTEM studies on the shape nc-Si/SiO_2 MLs [117, 118], as well as the results on nc-CdSe/SiO_x MLs [119], revealed an oval-like shape with the lateral dimension of nanocrystals exceeding the vertical dimension (see Fig. 2b). Transmission electron microscopy studies performed on Si/SiO_2 MLs have shown [117, 120] that nucleation starts in a-Si layers away from the interfaces. The growth of the crystalline nuclei is stopped by the interfaces, which accounts for the observed oval-shaped nanocrystals in ultra-thin nanocrystalline layers. A summary of the results on the relation between the initial thickness of amorphous layers and NC size indicates that the average vertical size of the crystals after conventional annealing is in good agreement with the thickness of the amorphous layers [63, 119, 121–123] for layers thinner than 7 nm. For thicker films, the nanocrystallite size is smaller than the thickness of the initial amorphous layers [117]. Though a quantitative analysis of the nanocrystallite size distribution was not obtained, it has been seen that the difference between recrystallized spheres was not greater than a few monolayers. Here we would like to pay attention to an observation of NCs whose size was larger than layer thickness. In case of ultra-thin SiO_x layers of Si/SiO_x a-MLs (a-Si layer thickness $d = 2$ nm), it has been noted [33] that Si crystallites were larger than 4 nm. The formation of crystallites larger than the initial thickness of layers has been explained with the presence of overstoichiometric silicon in the SiO_x layers, which made possible the observed growth of large Si NCs.

Crystallization Temperature Summarizing the results on thermal crystallization of a-Si layers in a-MLs, one can see an increase of the crystallization temperature by approximately 300 K when the Si layer thickness was reduced to 2 nm [45]. A similar behavior has been reported for a-Ge based MLs. The crystallization temperature of thick a-Ge film is about 770 K, while 1.7 nm a-Ge layers did not show crystallization even after annealing at 1030 K [124]. A different approach has been applied in order to study qualitatively the T_c changes of ultra-thin a-Se layers in Se/CdSe a-MLs. The a-Se crystallization was realized by exposing MLs with a laser beam which is strongly absorbed. High-frequency (vibrational) Raman spectra were measured [100]. It has been observed that a gradual increase of laser power density caused crystallization of Se sublayers; in similar low temperature (25 K) experiments (using even higher laser power densities), there was no indication of substantial crystallization, implying that the latter is mainly a thermal effect. From the threshold for crystallization, it was obtained that crystallization of three-dimensional a-Se occurs at laser beam power density twice lower than that necessary for crystallization of 3.5 nm layers.

Some results on the T_c increase of crystallization temperature with decreasing a-Ge and a-Si layer thickness are summarized in Figure 3. It is seen that decreasing the layer thickness below 10 nm results in a strong increase of T_c. For a-Si layers in Si/SiO_2 MLs this increase has been fitted by the relation [117]:

$$T_c = T_o + \Delta T \cdot \exp(-d/C) \tag{1}$$

where the thickness of the Si layers is d and the parameters are $T_o = 988$ K, $\Delta T = 663$ K, and $C = 2.43$ nm. The second term in Eq. (1) represents the shift of the crystallization temperature. There might be an additional influence of the composition of the "high-temperature" layers which is expected to have an influence on the constant C. The observed T_c increase has been connected with an existing built-in strain on a-MLs. The different T_c reported at equal Si layer thickness can be accounted for [117] with uncertainties in the layer thickness and the accuracy (±25 K) of the crystallization temperature. In addition, it has been assumed that a thinner "high-temperature" layer may not be able to release the same built-in strain in the ML structure which may change T_c [125].

A general discussion from thermodynamic point-of-view of the crystallization process in a-MLs consisting of alternatively deposited amorphous semiconductors ***A*** and ***B*** has been made in [59, 105]. Both the layer thickness and the nature of constituent materials (heterointerface) have been shown to be key factors for determining the crystallization temperature. According to the well-known theory of the

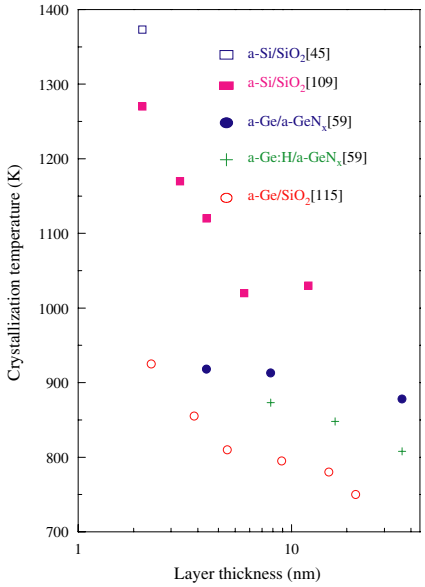


Figure 3. Crystallization temperature T_c of different Si- and Ge-based regular amorphous multilayers as a function of the layer thickness of the layers with lower T_c. A great T_c increase is seen in multilayers having layer thickness <10 nm.

homogeneous nucleation, crystallization of a total volume occurs only when the radius of crystalline nuclei (embryos) exceeds a critical value for the free energy of the system can be minimized. The crystallization temperature of an amorphous material is determined by the critical size $r_c = 2\gamma/\Delta G_v$, where γ and ΔG_v represent the interfacial free energy per unit area and the decrease of the bulk free energy per unit volume accompanied by the crystallization of super-cooled liquid. In the consideration of crystallization in a-MLs, the layer ***A*** was assumed to have a lower bulk crystallization temperature T_c than the material ***B*** and thickness smaller than the critical embryo radius r_c. For this reason, the crystallization nucleus formed in the material ***A*** were assumed, to a first approximation, to be cylindrical. Then, the free energy change upon crystallization is described as follows [66]:

$$\Delta G_{2D} = -\pi r^2 d\Delta G_v + 2\pi r^2 \Delta\gamma + 2\pi r d\gamma_{ac} \qquad (2)$$

where r is the radius of the cylindrical crystalline nucleus, ΔG_v is the decrement of the free energy per unit volume in amorphous-to-crystalline phase transformation, d is the layer thickness of the material ***A***, $\Delta\gamma = \gamma_{bc} - \gamma_{ba}$, is the increment of the interfacial free energy between material ***A*** and ***B*** per unit area when material ***A*** is crystallized, γ_{ac} is the interfacial free energy per unit area between amorphous and crystalline phase of the material ***A***, γ_{bc} is the interfacial free energy between the material ***B*** and the crystallized material ***A***, and γ_{ba} that between the material ***B*** and the initial amorphous material ***A***. ΔG_{2D} keeps increasing with r when $d < 2\Delta\gamma/\Delta G_v$, which means that crystallization never occurs. For $d > 2\Delta\gamma/\Delta G_v$

$$r_c = d\gamma_{ac}/(d\Delta G_v - 2\Delta\gamma) \quad \text{and}$$
$$\Delta G_{2D}^{\max} = \pi(d\gamma_{ac})^2/(d\Delta G_v - 2\Delta\gamma) \qquad (3)$$

Taking into account that $\Delta G_{2D}^{\max} \sim kT_c$ (k is Boltzmann's constant), the above relations show that T_c should increase with decreasing layer thickness for $\Delta\gamma > 0$, while T_c must decrease with decreasing d for $\Delta\gamma < 0$. It also predicts that not only the layer thickness but also the nature of the heterointerface affects T_c. The last prediction has been experimentally proven in Se/Se_xTe_{100-x} a-MLs. It has been shown [110] that crystallization temperature of a-Se layers in those MLs depends on the composition of Se_xTe_{100-x} layers. The results on a-Ge based MLs shown in Figure 3 are also in agreement with such a prediction.

Recently, a new model has been suggested [125] in which, in addition to the interface energies, the crystallization temperature of the thick amorphous layer, and the thickness of the layers, the melting point of the system was taken into account. In the framework of this consideration, the experimentally determined exponential increase of the crystallization temperature has been fitted by

$$T_c = T_{ac} + (T_{\text{melt}} - T_{ac})\exp(-d/C) \qquad (4)$$

where C has been found to be equal to 2.52–2.56 nm independent on the material (Si, Ge). In this model, it has been assumed that the crystallization nucleus is symmetrically embedded in the amorphous material between the interfaces and is cylindrical in shape. However, an additional spacing l has been introduced that corresponds to a finite separation of the nucleus from the boundaries represented by the second material ***B***. For $l = 0$, that is, the vertical nucleus size $h = d$, it has been assumed that a sharp interface is formed with the specific interface energy γ_{bc}. In the other limit of a very large l, the material ***B*** and nucleus are separated by two noninteracting well-defined interfaces, namely, the interface between the crystalline and amorphous phase of material ***A*** and the interface between the amorphous ***A*** and ***B***. For small l, the concept of an effective interface layer has been adopted and it has been assumed that a quasi-amorphous layer (called quasi-interface) is formed in the space between crystalline ***A*** and ***B***. The effective free interface energy γ_{bc}^{eff} is

$$\gamma_{bc}^{\text{eff}} = \gamma_{ac} + (\gamma_{bc} - \gamma_{ac})M \qquad (5)$$

where M is an effective order parameter which is normalized to unity for the real ***B***/crystalline ***A*** interface and zero for the amorphous/crystalline ***A*** interface. Under those assumptions, Eq. (2) becomes

$$\Delta G_{2D} = -\pi r^2 h\Delta G_v + 2\pi r^2 \gamma_{bc}^{\text{eff}} + 2\pi r h\gamma_{ac} \qquad (6)$$

and for the crystallization temperature, the following dependence has been obtained [125]:

$$T_c = T_{ac}\{1 + (\gamma_{bc} - \gamma_{ac} - \gamma_{ab})/\gamma_{ac}\exp(-d/4l_o)\} \qquad (7)$$

where T_{ac} is the crystallization temperature of a thick amorphous film ***A***, and l_o has been interpreted as an average screening or bonding length that is related to the range of interatomic forces typical for the crystalline ***A*** and ***B***. Through comparison with the empirical relation (4), which describes the experimental exponential increase of the crystallization temperature, the following relation has been derived [125]:

$$T_{\text{melt}} = T_{ac}(\gamma_{bc} - \gamma_{ab})/\gamma_{ac} \qquad (8)$$

According to the relations (4) and (7), the T_c increase in the limit $d \to 0$ is related to the difference in the specific interface energies of the interfaces between the crystalline and amorphous phases involved. The model yields a lower bound for the layer thickness below which no crystallization can occur in a-MLs. The value calculated for the Si/SiO_2 system is in good agreement with the experimental observation [125].

3.2.2. Laser-Beam Crystallization

Laser beam annealing of a-MLs under cw or pulse exposure has also been applied as an alternative of thermal crystallization, in order to produce nanocrystalline films [100, 103, 106, 109, 110, 121, 126–130]. In a series of studies on crystallization with an excimer laser of a-Si:H/a-SiN_x:H MLs [126–130], ability for fabrication of electroluminescence cells has been shown having visible light emission. This approach provides several advantages such as selected area processing, rapid crystallization, and avoidance of the surface contamination. It is also possible to crystallize a thin film while keeping the substrate temperature low enough to prevent its thermal damage.

The primary interaction between laser light and semiconductors is based on excitation of electrons and phonons [131]. The dominating interaction mechanisms depend on the laser parameters and the particular type of semiconductor, its microstructure (amorphous/crystalline), doping, etc. The interaction mechanisms may be thermal or non-thermal depending on the relative size of the thermal relaxation time and the time for structural rearrangements of material atoms or molecules. In most cases of semiconductor processing, the excitation energy is rapidly dissipated into heat (thermal laser processing previously mentioned). In spite of the thermal character, laser-driven processes may be quite different from those initiated by a conventional heat source. The laser-induced temperature rise can be localized in space and time. Temperature changes of more than 10^4 K can be induced in a small volume at a heat rating of 10^{15} K/s. For example, heat flow calculations have been performed for a-Si:H thin film for the case of single-shot laser annealing with an excimer laser [132]. It has been shown that the laser irradiation is able to melt and crystallize the film but even the melting process could not be described by pure heat flow calculation. It is interesting to note that the results of crystallization of a-Si:H thin films with a cw Ar^+ laser and a pilsed Nd:YAG laser have indicated [133] that the mechanisms of crystallization are different and, in particular, the pulsed crystallization is a nonthermal process. Another example of the complex interaction of the light with matter is laser-induced crystallization of a-Se. Polarization-dependent photocrystallization has been observed and though it seems to be optical in origin [134], an alternative, essentially thermal, mechanism has also been discussed [135].

Both thermal and pure photo-induced crystallization has been observed in a-MLs. Compositional and thickness dependent changes of the crystallization process in amorphous $Se/Se_{100-x}Te_x$ multilayers ($x = 7.5$, 15, and 30) have also been studied under light treatments [109, 110]. The treatments have been performed using the 647.1 and 530.1 nm Kr^+ laser lines. The thermal effect of the laser exposure has been investigated at room temperature in $Se/Se_{70}Te_{30}$ a-ML, using the red laser line (which is strongly absorbed by the $Se_{70}Te_{30}$ sublayers but weakly so in the Se ones), as well as on $Se/Se_{100-x}Te_x$ a-MLs using the green laser line [110]. In these experiments, it has been established that the threshold power for crystallization (and, hence, the crystallization temperature) of $Se_{70}Te_{30}$ sublayers increases with decreasing their sublayer thickness. The result is in accordance with those reported previously for a-Se [100, 106].

"Pure" photo-induced structural changes in $Se/Se_{100-x}Te_x$ a-MLs have been explored at a low temperature (38 K) at which thermally induced photo crystallization is not anticipated. It has been found [110] that exposure to the green laser line lead to an increase of the strength of photo-induced crystallization in $Se/Se_{100-x}Te_x$ a-MLs with increasing Te content, which is in agreement with previous reports on three-dimensional $Se_{100-x}Te_x$ alloys [136–138]. This result implied that there is no considerable change in Se-Se and Se-Te bond energies when a dimensionality reduction (from 3 to 2) takes place in the alloys. Furthermore, the thickness dependence of the strength of photo-induced crystallization was studied in $Se/Se_{70}Te_{30}$ a-MLs for which the strongest effect has been observed. A series of Raman scattering spectra measured with the laser beam focused on the same spot of each sample are shown in Figure 4. One can see that the intensity of the 237 cm^{-1} band, which corresponds to scattering from trygonal Se, and hence, the relative part of the crystalline phase in both $Se_{70}Te_{30}$ and Se layers increases when the time of preliminary exposure increases. A reduction of overall Raman intensity has also been observed. The latter has been ascribed to a deterioration of the ML structure and/or the surface quality of the sample due to crystallization. A conclusion has been derived [109, 110] that, as in the case of thermally induced crystallization, the rate of photo crystallization in $Se_{70}Te_{30}$decreases with decreasing sublayer thickness but the cause of this dependence in the case of photo crystallization is different from that responsible for thermal crystallization. The last suggestion is based on the mechanism of photo-induced crystallization developed for a-Se [134]. According to this model, the photo crystallization strength in a-Se is controlled by the production of electron hole pairs and not by the density of absorbed energy. This implies that structural disorder may play a significant role in the process of photo crystallization. Thus, the observed lower rate of crystallization in $Se/Se_{70}Te_{30}$ a-ML having thinner sublayers has been attributed [110] to a decrease in the lifetime and, hence, the concentration of photo-excited free carriers due to a disorder level increase. Moreover, in thinner sublayers, photo crystallization starts at a higher level of structural disorder and, hence, at the same light exposure, a longer time is necessary in order to achieve the same crystallization effect.

Laser annealing along with both previously described thermal post-treatment methods (furnace annealing and rapid thermal annealing) were applied for crystallization of a series of a-Si:H/a-SiN_x:H MLs [139]. It has been shown that when the thickness of a-Si:H layers is larger than 4.0 nm, the constrained a-Si:H layers were well crystallized by each of three methods. However, when the thickness of a-Si:H layers is smaller than 4.0 nm, the laser annealing method is the most advantageous for the formation of the nuclei and

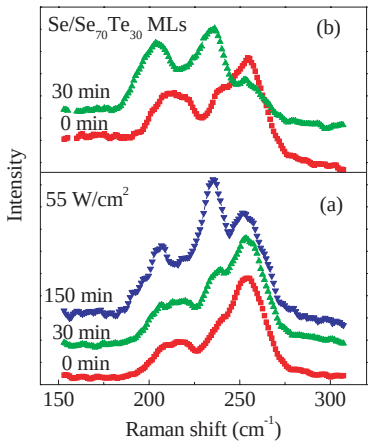


Figure 4. Raman scattering spectra of Se(4 nm)/$Se_{70}Te_{30}$ MLs having $Se_{70}Te_{30}$ layer thickness of (a) 3.5 nm and (b) 7.0 nm. The spectra were measured on a constant spot for each sample after a preliminary laser beam exposure (647.1 nm, 55 W/cm^2) for various times indicated in the figure.

crystallization of the samples compared to the other methods. Using laser annealing, ultra-thin a-Si:H layers can be well crystallized even when their thickness decreases down to 2.0 nm. In contrast, the nucleation become more difficult upon furnace annealing, so that few Si NCs can be formed by this method.

3.2.3. Crystallization Induced Changes

The crystallization of layers in a-MLs by annealing or laser exposure can cause a number of structural and mechanical changes. Among them are cracking [103], some increase in the magnitude of the internal strains [123], deterioration of the interface smoothness, ML periodicity [105, 140], ML surface [110], etc.

Laser power between 0.25 W and 3.0 W from an Ar^+ cw laser were used for crystallization of A-Ge:H/a-SiN_x:H MLs [103]. No visible changes or measurable signs of crystallization have been detected for MLs on crystalline silicon substrates. For multilayers on glass, however, changes have been observed. Even for powers below 1.0 W, some regions appeared "milky," caused by microscopic bubles that increased in density with power. For powers greater than 1 W, some rough and shiny areas appeared, corresponding to severe cracking and pitting when viewed through a microscope. The appearance of such mechanical defects can be avoided using slow thermal annealing [103].

As previously mentioned, in order to determine the average size of nanocrystals, X-ray diffraction and HRTEM studies are usually used. The Scherrer's equation

$$d_{NC} = \lambda/\delta 2\Theta \cdot \cos\Theta \tag{9}$$

is employed to a distinct band in the X-ray diffraction pattern of the sample in order to calculate the average nanocrystallite diameter d_{NC} (Fig. 5). Here, Θ and $\delta 2\Theta$ are the position and full width at half the maximum of the band. However, a discrepancy has been reported [119] between the average size of CdSe nanocrystals in nc-CdSe/SiO_x MLs determined with X-ray diffraction and HRTEM studies. It is known that the full width at the half maximum of X-ray diffraction peaks depends not only on the NC size but also on existing microstrains and deformations in the NC network. The observed discrepancy has been related to the existence of microstrains in the nanocrystals. Using relation [141]

$$(\delta 2\Theta \cos\Theta/\lambda)^2 = 1/d_{NC}^2 + 16\varepsilon^2(\sin\Theta/\lambda)^2 \tag{10}$$

and the average nanocrystallite diameter obtained from HRTEM, the approximate levels of the microstrains have been estimated assuming the CdSe NC size and strain (ε) distributions to be Gaussian. Values of around $20 \cdot 10^{-3}$ and $28 \cdot 10^{-3}$ have been obtained for nanocrystals having $d_{NC} =$ 10.0 and 5.0 nm, respectively. They are close to the values of $20–22 \cdot 10^{-3}$ calculated [142] for the CdSe layers of a free-standing CdS/CdSe superlattices with layers of equal thickness. A value of $\varepsilon \approx 49 \cdot 10^{-3}$ has been determined for nanocrystals having $d_{NC} = 3.0$ nm, which indicates that, as it might be expected, the level of microstrains rises with decreasing NC size. Several reasons can cause this strain. The surface-to-volume ratio increases with decreasing layer thickness and nanocrystallite size, which is correlated with a higher intrinsic stain and caused by bond-angle deviation. Some interface strain can originate from the difference in the thermal expansion coefficient of the two materials. There may be also an intrinsic component resulting from the volume shrinkage with the transformation of an amorphous region into a more dense crystalline material. Keeping in mind that both the particle size and the lattice strain have a different angle dependent influence on the Bragg diffraction linewidth, the particle size and strain can be separated by two-line Scherrer analysis as described in [143].

The internal strain distribution has also been studied in nc-Si:H/a-Si:H MLs [144] by analyzing the intensity and line widths of X-ray diffraction peaks in the low- and high-angle ranges and the obtained results have been compared with those from Raman scattering. The elastic tensile strain ε has been calculated approximately by using the X-ray diffraction angle of the (120) plane for Si single crystal ($2\Theta = 28.47°$), which was reduced to $2\Theta \approx 27°$ in MLs [145]. The relation

$$\varepsilon = (a - a_o)/a_o \tag{11}$$

has been applied, where a_o is the lattice constant of single crystal Si and a is the measured lattice constant of MLs due to the internal strain. The elastic tensile strain has also been estimated from the Raman scattering data using the relation

$$(\omega_c - \omega_s)/\omega_c = -\varepsilon(P + 2Q)/\omega_c^2 \tag{12}$$

where $P = 1.43\omega_c^2$ and $Q = (-1.89)\omega_c^2$ are the phonon deformation potentials of Si, ω_c is the frequency of the crystalline peak, and ω_s is the vibrational frequency of the shell region between the crystalline core and outer space in nanocrystallites [123, 146]. The effective strains obtained by both approached are about 5% for MLs with nc-Si:H layer thicknesses of 2.2 and 3.0 nm and a-Si:H thickness of 8 nm. The strain can partially be released by using high annealing temperatures [118]. A quasi-equilibrium annealing (with a rate of temperature increase of ~10 K/min) from 870 K to 1320 K has been proposed [63] to reduce internal strains and

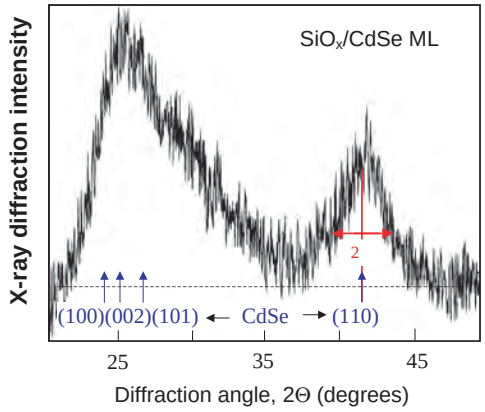


Figure 5. X-ray diffraction spectrum of a SiO_x(5 nm)/CdSe(5 nm) ML. The (100), (002), and (101) bands of wurtzite CdSe are not resolved while the (110) one is well resolved and can be used for determination of CdSe nanocrystallite size by applying Eq. (1); $\delta\Theta$ indicates full width at half maximum of the (110) band.

interface defect density upon rapid thermal crystallization of Si/SiO_2 a-MLs.

Small- and high-angle X-ray diffraction studies on a-Se/a-CdSe MLs, annealed at various temperatures for a long time (18 hr) have shown [105, 140] that the ML periodic structure vanishes upon annealing at 333 K (the glass transition temperature of a-Se is $T_g \approx 320$ K). Upon annealing at 363 K, a crystalline phase was evident by its main diffraction peak situated at $2\Theta = 29.6°$. Also, upon annealing at 720 K in a vacuum used for preparation of a-Si:H/nc-Si MLs [147], it has been seen that the layered structure is partially destroyed. This observation has been explained by the effusion and redistribution of hydrogen over the volume of the films. On the other hand, the SAXRD spectra of a-Ge:H/a-GeN_x MLs annealed at temperatures higher than the crystallization temperature of a-Ge (723 K) have proved that both the ML period and the interface quality did not change appreciably after the crystallization of a-Ge:H layers [101]. The HRTEM results on those MLs, as well as on other Ge-based a-MLs [67, 103], have displayed that a-Ge layers crystallized without disrupting the multilayer structure, and that after crystallization the interfaces were atomically smooth and uniform. It has also been established that in a-Si:H/SiO_2, a-Si:H/SiC_x, and a-Si:H/SiN_x [63, 121], as well as in SiO_x/GeO_y [148] MLs, both thermal and laser-beam crystallization of the a-Si:H and GeO_y layers do not destroy the regular periodic structure of multilayers. Finally, SAXRD studies of CdSe/a-SiO_x MLs, annealed at 673 K for 60 min [93], have shown that these MLs also exhibited a very good periodicity. The above short review of the results on the periodicity of various nanocrystalline/amorphous multilayers shows that the periodic structure of amorphous multilayers is resistant to annealing or laser beam illumination when the constituent material with the higher crystallization temperature has a rigid structure. In such MLs, the growth of semiconductor microcrystals in the "low-temperature" sublayers is restricted along the ML axes and, thus, crystallization does not destroy the ML periodic structure.

3.2.4. Material Intermixing

The furnace crystallization of sublayers from the "low-temperature" material in a-Ge and a-Si based a-MLs, requires quite high annealing temperatures (>770 K for a-Ge:H [1] and >970 K for a-Si:H [23]). At these temperatures, undesired interdiffusion and alloying of the constituent materials may take place at ML interfaces. In certain studies on annealing of a-Ge:H/a-GeN_x [101], a-Ge:H/a-SiN_x:H [103] and a-Ge/a-C [67] MLs alloying has not been observed. On the other hand, in a-Si:H/a-Ge:H MLs [111], the annealing at temperatures > 670 K leads to strong material intermixing, followed by the crystallization of the resulting silicon germanium alloy. Anisotropic Ge diffusion has been directly observed in SiGe/Si-strained superlattices [149] induced by nonuniform strain in SiGe/Si interface in the range 700–1000 °C. Additionally, it has been shown [150] that the Ge-Si interdiffusion increases by more than two orders of magnitude for stucked hut clusters and the degree of Si alloying in vertically aligned self-assembled islands increases with the number of stached layers. Interdiffusion has also been reported in annealed Si/SiC [112] MLs. A small region of mixed Ge-Si alloy at each interface has been observed in SiO_x/GeO_y MLs annealed 770 K [148].

Although thermal crystallization in chalcogenide a-MLs occurs at much lower temperatures than in a-Si- and a-Ge-based MLs, material intermixing has also been observed in MLs of various materials. The effect of annealing (at 340 K for 60 min) on photo and electrical conductivity, as well as on infrared absorption of As_6Se_{94}/Te, a-MLs has been studied in [70]. It has been observed that at the first step of the annealing process, a diffusion of Se atoms into Te sublayers occurred and As-Se/Se-Te a-ML formation took place. Further annealing leads to As-Se/As-Se-Te MLs formation. Systematic optical and electrical studies on photo- and thermo-induced changes of As_2S_3/Se, As_2S_3/Se_xTe_{1-x} ($0.1 < x < 0.5$), and $As_6Se_{94}/Se_{80}Te_{20}$ a-MLs have indicated [80, 81] that the periodic structure of these MLs was changed by thermal treatment and laser beam exposure due to the interdiffusion. It should be pointed out that in this group of MLs, the corresponding photo-induced changes of their optical parameters (absorption, refraction) can be successfully used for optical recording. Raman scattering measurements have been carried out on as-deposited and annealed at 670 K multilayers of ZnSe/CdSe [151]. It has been shown that $Zn_xCd_{1-x}Se$ interface regions were formed during sample preparation and a gradual change in the composition is characteristic for these regions. Raman spectra of ZnSe/CdSe MLs annealed at 670 K for 60 min are displayed in Figure 6. It can be seen that the intensity of the band at around 225 cm^{-1} (it is due to scattering from $Zn_xCd_{1-x}Se$) is rather high. This indicates that large $Zn_xCd_{1-x}Se$ interface regions exist in annealed samples. The result has been explained assuming diffusion of Cd atoms from CdSe into ZnSe. It has been shown in this study that when the nominal thickness of CdSe layers is very small, taking advantage of material alloying at the interfaces, it is possible to produce $Zn_XCd_{1-X}Se$ nanoparticles in a ZnSe matrix by means of thermal evaporation of CdSe and ZnSe in vacuum.

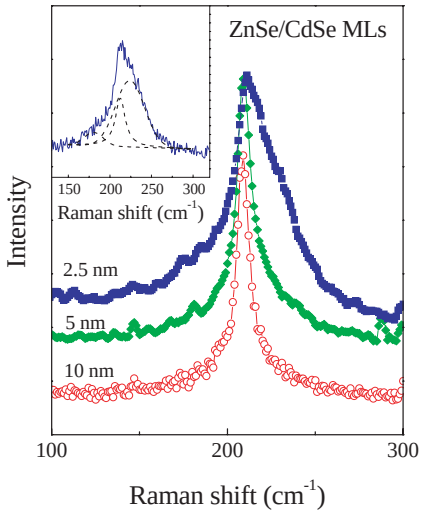


Figure 6. Raman scattering spectra of ZnSe/CdSe MLs with layer thicknesses indicated in the figure. The 647.1 nm line of a Kr^+ laser was used for the excitation. Inset: Raman spectrum (solid line) fitted to one Lorentzian peaked at 212 cm^{-1} and two Gaussians at 180 cm^{-1} and 225 cm^{-1} (dashed lines).

Various results obtained under annealing of different amorphous MLs may be understood if one keeps in mind that several competitive factors can affect material intermixing and alloying at the ML interfaces. Diffusion coefficients of the atoms of each constituent material in the other one, and the possibility for alloying of both materials (i.e., the correlation between the magnitudes of the bond energies in the constituent materials and their alloy) could be among them. Hence, various approaches may be applied in order to avoid the layer intermixing. One of them is the selection of proper pairs of materials; the data reviewed imply that semiconductor/semiconductor oxide (or nitride) pairs are better than semiconductor/semiconductor ones. This suggestion is motivated by the results for nc-CdSe/SiO_x [152], nc-Ge/a-GeN_x [101], and nc-Ge/a-SiN_x [103] MLs, in which no alloying has been observed and the considerable alloying reported for nc-CdSe/ZnSe, nc-CdSe/GeS_2 [151, 153], nc-Ge/a-Si:H [111], etc. The second approach, which may be applied to avoid thermally induced sublayer intermixing, is the application of combined photo-thermal or "pure" photo-induced crystallization processes. As previously described [79, 86], a considerable diffusion of Se and Te takes place in chalcogenide materials upon long-time furnace annealing even at considerably low temperatures. In Se-based MLs, one may use the fact that "pure" amorphous selenium and, particularly, selenium-tellurium alloys exhibit strong photo thermally-induced crystallization at room temperature. Moreover, they crystallize under laser beam exposure at fairly low temperatures (<100 K), at which the thermally induced atomic diffusion is extremely slow. Thus, using photo crystallization at room and low temperatures, it is possible to make optical recording in amorphous chalcogenide MLs with minimum alloying at the ML interfaces. It has been shown [103] that crystallization of a-Ge in a-Ge:H/a-SiN_x:H MLs is a type of explosive process, caused by the high stress of the a-Ge:H, which is triggered by laser heating. In this situation, crystallization occurs for a very short time and no interface alloying has been detected. All given examples indicate that special attention should be paid to material intermixing, in particular, when furnace long-time thermal crystallization of a-MLs is applied, As for the short-time crystallization procedures (rapid thermal annealing and laser annealing), it seems that they do not deteriorate the interface quality and ML periodicity very much.

All silicon and germanium-based multilayers described in Section 3 are listed in Table 1 along with the techniques applied for their preparation and crystallization of the layers with the lower T_c. This information for chalcogenide-based MLs considered in Section 3 is summarized in Table 2.

4. DISCONTINUOUS NANOPARTICLE LAYERS

4.1. Nanostructured Composites

Fabrication of semiconductor nanoparticles is very important for realization of new functional quantum devices [154]. Various techniques, such as doping in glasses, epitaxial nanocrystal growth, gas evaporation, cosputtering, laser ablation, colloidal methods, ion implantation, etc. have been applied [144 and references therein] for fabrication of nanoparticles. They can be produced as nanophase powders or nanostructured composites. The latter have two main advantages to the nanophase powders: they ensure a good mechanical stability and keep nanoparticles from the influence of the surrounding ambient. Bulk composites of $CdSe_xS_{1-x}$ semiconductors and silicate glasses (so-called semiconductor-doped glasses) have been the first to focus much interest with their high-optical nonlinearity [156–158] and quantum-size changes of the electronic structure [32, 35]. Thin films of CdS [152, 159, 160], CdSe, and CdTe [161] dispersed in high-frequency sputtered SiO_2 have also been prepared as counterparts of the doped glasses. In those bulk samples and thin films II–VI, semiconductor nanocrystals have been grown in a diffusion-controlled phase decomposition of oversaturated solid solution. That process is carried out at high temperatures ($T > 900$ K) for several to 100 hr [159]. Silicon nanoparticles in Si-rich silicon oxide films are also produced by high-temperature annealing ($T > 1170$ K) in an inert atmosphere [162–170]. In this approach, the average nanoparticle size could be tuned either by varying the annealing temperature or by changing the excess silicon content in the deposited films. Recent investigations have shown that very high photo conductivity can be obtained in hydrogenated amorphous silicon films containing a small volume fraction of nanocrystalline silicon inclusions [171]. However, even playing with the deposition parameters, it is rather difficult to control the size, amount, and spatial distribution of those nanocrystallites.

Systems based on metallic clusters encapsulated in an insulating or semiconductor matrix have also attracted considerable interest with their interesting optical properties [172], as well as with the appearance of switching effects in nanodisperse metals [173–175]. A somehow similar effect has been intensely investigated in metal/a-Si:H/metal sandwich structures [176–178] and MLs having discontinuous metal layers [179]. Multilayer systems of amorphous silicon and metals have been used as mirrors for soft X-ray radiation [180], X-ray detectors [181], and position sensors [60]. It has also been demonstrated that a-Si/Mo and a-Si/Ta MLs exhibit very interesting features such as linear lateral photo voltage [182] and giant current fluctuations [179] when the metal layers are discontinuous. Similarly, to the fabrication of semiconductor-doped glasses, the common method to prepare materials consisting of nano-sized metallic clusters embedded in insulating matrix is cosputtering or coevaporation of the metal and insulator on a suitable surface. The cluster sizes are controlled by varying the substrate temperature and/or post-growth annealing.

4.2. Preparation of Discontinuous Layers

An alternative approach for preparation of nanostructured composites is to use sequential vapor deposition of semiconductors (metals) and insulating (semiconductor) materials. The Stranski-Krastanow self-organized epitaxial growth of III–V [183] and, more recently, II–VI semiconductor nanocrystals [184, 185] in strained systems has attracted much attention. This mechanism takes place on the smooth surface of a crystalline material, but it can explain the nanocrystal growth observed on a rough surface [180, 185–189]. This nonepitaxial self-organized nanocluster formation has

Table 1. Silicon and germanium-based regular multilayers.

Multilayer	Preparation technique	Crystallization
a-Si:H/a-SiN_x:H	plasma enhanced CVD [139]; radio frequency plasma enhanced CVD [126–130];	furnace annealing [112]; rapid thermal annealing [139]; laser beam crystallization [126–130];
a-Si:H/$Si_{1-x}C_x$	glow discharge	furnace annealing [112], interdiffusion
a-Si:H/a-Ge:H	glow discharge	furnace annealing [62, 116], interdiffusion
a-Si/SiO_2	plasma enhanced CVD [61, 63, 64, 67]; rf magnetron sputtering [63, 118, 122]; low pressure CVD [90, 91];	furnace annealing [63, 64, 90, 91, 124]; rapid thermal annealing [125];
a-Si:H/SiO_2	low pressure CVD, layer-by-layer plasma oxidation	rapid thermal annealing plus furnace annealing [56]
a-Si/SiO_2	electron beam evaporation, *in-situ* plasma oxidation; magnetron sputtering [70]	furnace annealing [48–50, 61, 67]
a-Si/SiO_2	molecular beam epitaxy [45–47]	
nc-Si/SiO_2	rapid thermal CVD	[57]
nc-Si/a-Si	glow discharge	intermediate annealing in a hydrogen plasma [147]
a-Si:H/a-Ge:H	glow discharge	furnace annealing [62, 116], interdiffusion
a-Ge/SiO_x		furnace annealing, interdiffusion [124]
a-Ge:H/a-SiN_x:H	rf glow discharge CVD	furnace annealing [103]
a-Ge:H/a-GeN_x	rf reactive sputtering	furnace annealing [66, 101, 114]
a-Ge/a-GeN_x	rf reactive sputtering	furnace annealing [66, 101, 114]
a-Ge/a-Se		furnace annealing [74]
a-Ge/a-C	magnetron sputtering [14]	
a-Ge_xC_y:H/a-Ge_xC_y:H	continuous plasma enhanced CVD [65]	
SiO_x/GeO_y	reactive sputtering of Si and Ge	furnace annealing [148]

been used for fabrication of nanoparticle layers buried in thin film matrices. It occurs on the rough surface of amorphous, thin films deposited by means of physical evaporation, sputtering, or glow-discharge techniques.

4.2.1. Metal Nanoparticle Layers

Metal (Mo,Ti,V)/a-Si ML systems (also called composite films) have been prepared by magnetron sputtering of both materials or by an rf glow discharge deposition of a-Si and thermal evaporation of metals [180, 186]. Discontinuous metal sublayers were formed by interrupting the growth process before the stadium of coalescence at a nominal layer thickness of around 1 nm. The thickness of a-Si layers was 100 nm. Scanning force microscopy measurements of surface topography revealed metal clusters with diameters of the a-Si:H surface roughness (2–8 nm). These clusters were arranged in chains up to several hundreds of nanometers in length. Cross-section TEM micrographs showed island-type metal layers having a typical thickness of 10 nm. It has been assumed that metal clustering is supported by the existing valleys of the a-Si:H surface structures. As in MLs with continuous layer, an material mixing has been observed at the interface. It is interesting to mention the observed asymmetry of the mixing. The mixed interface region from the growth of the metal on Si has significantly greater thickness ($\sim$10 nm) than that (a few nanometers) from the growth of the Si on the metal.

Nanoparticle layers of Ga, with relative low-size dispersion ($\leq$20%) and regular shape of a truncated sphere, have been produced by thermal evaporation of Ga on 5 nm thick SiO_x films. Metal particles were formed in the liquid state by a self-organization process [189, 190], which is related to the partial wetting character of Ga with respect to the insulator [190]. The melting point of Ga in nanoparticle layers is considerably lowered with respect to the bulk materials and it is liquid at room temperature [191].

Self-organized growth has also been observed on deposition of cobalt on alumina [192–195]. Multilayers were

Table 2. Chalcogenide-based regular multilayers.

Multilayer	Preparation technique	Crystallization
a-Se/a-$Se_{100-x}Te_x$ ($7.5 \leq x \leq 30$)	thermal evaporation [109, 110]; laser beam evaporation [82–85, 93]	laser beam crystallization [109, 110]
a-Se/CdSe	thermal evaporation [76, 77, 88, 99, 100, 105, 140]	furnace annealing [105, 140]; laser beam crystallization [100]
a-$Se_{85}Te_{15}$/CdSe	thermal evaporation [24, 27, 30, 76, 78]	
a-Se/As_2S_3	thermal evaporation	laser beam crystallization [106]
a-Se/As_2Se_3	thermal evaporation	furnace annealing [72]
a-Te/As_6Se_{94}	thermal evaporation	furnace annealing [79], interdiffusion
As_2S_3/Se_xTe_{100-x} ($10 < x < 50$)	thermal evaporation	laser beam processing [80], interdiffusion
As_2S_3/AsSe	thermal evaporation	laser beam processing [81], interdiffusion
nc-CdSe/a-SiO_x	thermal evaporation	furnace annealing [93–96, 117, 201]
nc-CdSe/a-GeS_2	thermal evaporation	furnace annealing [97, 98, 201], interdiffusion
nc-CdSe/ZnSe	thermal evaporation	furnace annealing [97, 98, 151], interdiffusion

prepared by sequential rf sputtering of Co and Al_2O_3 at room substrate temperature. It has been shown [193, 194] that nearly spherical crystalline Co clusters with diameters ranging from 0.8 nm to 4.5 nm were formed when nominal deposited Co thickness increases from 0.2 to 1 nm. The Al_2O_3 layer thickness was several nanometers. The metal nanocluster formation has been discussed in terms of a large difference of surface energies of the transition (and noble) metals and insulators [196]. Additionly, cross-section TEM of Co(0.7 nm)/Al_2O_3(3 nm) MLs has shown [195] an ordered structure (seen in a given range of thicknesses), which consists of Co clusters with a diameter of 3 nm, whose vertical arrangement from plane to the plane is not random but shows an *additional* self-organization related to topology reasons [195]. When Co is deposited on Al_2O_3, clusters are formed that display the in-plane order. The next Al_2O_3 layer is assumed to wet perfectly this Co island-type layer being undulated. During the next Co deposition, cluster formation takes place preferentially in the hollows of the wavy Al_2O_3 surface. The periodic repetition of this phenomenon is considered to be the origin of the self-organized Co growth observed. The last result illustrates that in contrast to a codeposition method, the multilayer approach allows preparation of composite structures in which the size and spacing of the clusters can be controlled directly by varying the thickness of the layers. If this fabrication method can be extended to ferromagnetic alloys, it might open a new route for the fabrication of advanced magnetic recording media.

As previously discussed, at semiconductor/semiconductor interface, significant material intermixing may take place which can strongly effect ML properties. Similarly, in a metal/oxide structure, one should pay attention to possible reactions at the interface. X-ray photoelectron spectroscopy has been used to probe metal (Au, Ag, Mg, Mg-Ag alloy)/AlO_3 interface structures. It has been found [197] that there is no chemical reaction at (Au, Ag)/AlO_3 interfaces, while there is a significant reaction/diffusion at (Mg, Mg-Ag alloy)/AlO_3 ones. The reaction involves formation of Mg and metallic Al and then followed by metallic Mg diffusion to the interface.

4.2.2. Semiconductor Nanoparticle Layers

Preparation Semiconductor nanocrystals (quantum dots) of ZnCdSe has been grown by molecular beam epitaxy on a ZnSe crystalline surface which was intentionally roughened [198]. High-resolution transmission electron microscopy studies have shown that the vertical size of nanoparticles was significantly smaller than their in-plane size and, also, the size distribution of those dots was rather wide.

Another technique has been developed [97, 188, 199–201] for fabrication of CdSe nanoparticle layers buried in amorphous thin film matrices. Consecutive deposition of ultra-thin films of thermally evaporated CdSe on a relatively rough surface of the matrix material (SiO_x, GeS_2, ZnSe) has been applied. When developing this technique, it has been taken into account that during the nonepitaxial growth of thin films, the top surface of these films is rough with amplitude of this roughness varying with the deposition techniques. In case of thermal evaporation of materials, this amplitude is $\sim d^{1/2}$ (d-film thickness) [202] and, hence, the thicker the thermally evaporated layer, the greater its surface roughness. Multilayer of consecutively deposited SiO_x (GeS_2, ZnSe) and CdSe layers were fabricated in which the SiO_x thickness d_{SiO_x} was 20 times greater than that of CdSe layers d_{CdSe}. One can expect that CdSe layers will be discontinuous rather than continuous. A schematic diagram of such a ML structure is shown in Figure 7. Cross-section high-resolution electron microscopy measurements have revealed [97, 199, 200] that, indeed, CdSe does not form a continuous layer. Instead, nearly spherical CdSe nanoparticles partly isolated, and partly in contact among each other, are disposed in a composite SiO_x-CdSe "sublayer" of up to 10 nm in thickness (Fig. 8). It should be pointed out that nanoparticles are not grown in the surface valleys but their spatial distribution follows the morphology of the surface underneath. The great width of the CdSe nanoparticle spatial distribution also indicates that they should not be formed only in the surface valleys. The results indicate that the surface roughness favors the self-organized formation of CdSe nanoparticle layers rather than the growth of continuous thin films.

High-density Si nanoparticles have been successfully grown on Si_3N_4 and SiO_2 thin film surfaces by hot-wire CVD using disilane, in which Si atoms were generated on a heated tungsten filament [203]. A highest density of 1.1×10^{12} cm^{-2} has been achieved with corresponding average nanoparticle size of about 5 nm. Sandwich-structured thin films of silicon nanoparticles embedded in Al_2O_3 matrices have been prepared by pulsed laser beam evaporation [203]. It has been found that the sharpness of the Si/Al_2O_3 interface depend strongly on annealing conditions.

The described approaches of production of nanoparticle layers buried in a thin film amorphous matrix makes possible the reproducible fabrication of nanocrystals for shorter times and at considerably lower annealing temperatures than those usually used for the nanocrystal growth in

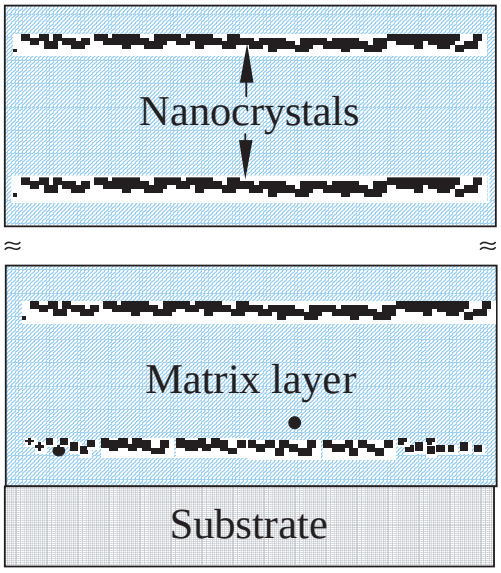


Figure 7. Schematic presentation of a semiconductor-matrix composite film in which the nominal semiconductor layer thickness d_s varies between 1 and 6 nm and the matrix layer thickness is $d_m = 20\ d_s$.

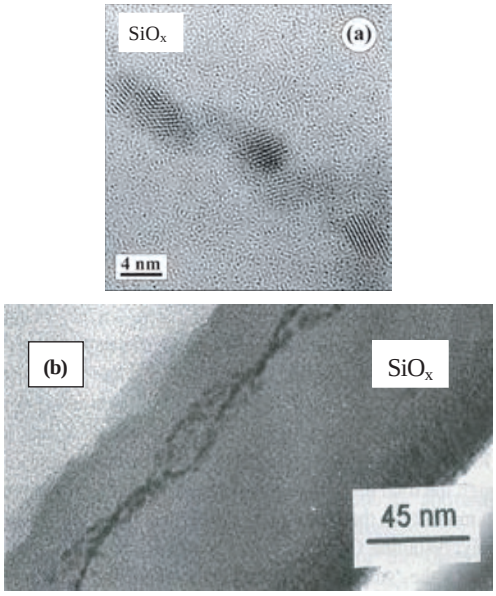


Figure 8. Cross-section view of the spatial arrangement of CdSe nanoparticles in as-deposited three layer structures SiO_x(40 nm)/CdSe(2 nm)/SiO_x(20 nm) at high (a) and low (b) magnification. The spatial distribution of the nanoparticles follows the surface morphology of the SiO_x layer.

semiconductor-doped glasses and their thin film counterparts. This technique, if applicable, can be very useful, especially for semiconductors such as Se, which form nanocrystals of very different sizes in the standard procedure of crystal growth in a diffusion-controlled phase decomposition of oversaturated solid solution.

Size Distribution and Structure A deeper inspection of the buried CdSe nanoparticles by high-resolution electron microscopy have shown [199] that nanoparticle crystal lattice is randomly oriented. Occasionally, deviations from the spherical shape have been observed indicating the effect of particle interaction by coalescence. The average size of CdSe NCs was estimated from the (119) band in the X-ray diffraction patterns of samples with various nominal CdSe sublayer thicknesses employing Scherrer's Eq. (9) and the values obtained were close to the nominal thickness of deposited CdSe layers. On the other hand, HREM evaluation of the particle sizes yielded a mean diameter, which is more than twice greater than the nominal layer thickness of deposited CdSe. Keeping in mind the above discussed reason for this discrepancy (internal strains and deformations), a value of $\varepsilon \approx 49 \cdot 10^{-3}$ has been determined for nanocrystals having $d_{NC} = 3.0$ nm, which indicates that as it might be expected, the level of microstrains rises with decreasing NC size.

The changes upon annealing of the SiO_x-CdSe samples at 670 K have been studied by a cross-section HRTEM [199, 204]. It has been observed that after annealing at 670 K, crystal lattice of the CdSe nanoparticles remains oriented randomly, but their sizes increase a little and the spherical appearance is improved. From the high-resolution imaging of lattice plane fringes and by comparison of the diffractograms of both images, it has been estimated that the crystallinity of the CdSe nanoparticles is improved upon annealing. Values of 0.12 and 0.15 have been obtained for η/d_{NC} ratio in as-deposited and annealed SiO_x(20 nm)/CdSe(1 nm)/SiO_x(20 nm) structures, respectively, (where η is half the width at half the maximum of the size distribution). In annealed SiO_x(40 nm)/CdSe(2 nm)/SiO_x(20 nm), $\eta/d_{NC} \approx 0.21$. This nanocrystallite size distribution is compatible and even slightly narrower than those reported for CdS [205] and Ge [206] nanocrystals grown in SiO_2 matrix during the annealing of thin films of CdS-doped silica and $Si_{1-x}O_xGe_y$, respectively.

One important observation from the HRTEM micrographs is the high density of planar lattice defects in both as-deposited and annealed films. These defects could be among the reasons for the great difference in the average nanocrystal sizes obtained from the X-ray diffraction patterns and those seen in HRTEM micrographs. This raises the question for the lattice type present in the nanoparticles. Usually, CdSe is described by the hexagonal wurtzite B4 type, which was found also for CdSe nanoparticles from HRTEM studies [207, 208]. On the other hand, it is known that in the nanoscale range II–VI, semiconductors tend to change rather easily from the wurtzite to the cubic zincblende B3 type and vice versa [209, 210]. This transition and the corresponding stacking disorder have been rated by measuring the lattice fringe spacings of a number of particles [204]. It has been observed that in as-prepared samples (deposited at room substrate temperature), CdSe particles display a predominantly wurtzite lattice type, while in the annealed structures (at 670 K), CdSe particles of the sphalerite lattice type have been frequently seen. For the as-deposited nanoparticle layer, three-quarters of the particles reveal spacings of the hexagonal and one-quarter spacings of the cubic or both lattice types. For the annealed structures, however, it has been found that approximately one-half of the particles reveal spacings of the cubic and another half spacings of the hexagonal or both lattice types, respectively. It is important to notice that in three-dimensional CdSe films deposited, the situation is just the opposite. The films deposited at low substrate temperatures ($T < 370$ K) show predominantly sphalerite structure. The annealing of three-dimensional films at 670 K converts their structure to wurtzite type. This difference in the behavior of three-dimensional and zero-dimensional CdSe is most likely due to surface-induced changes in bond strength in the nanoparticles.

Mechanism of Semiconductor Nanoparticle Formation As discussed in Section 2, at the first stage of thin film deposition, the formation of embryos takes place, which depends on the temperature, chemical nature, structure, and clearness of the substrate surface [202]. In particular, the presence of steps on the substrate surface induces an increase in the embryo concentration at these steps. It has also been shown [211] that the curvature and stress at a rough or even disordered surface, strongly affect reaction rates at solid interface as they create different environments at different reaction sites. The sequential vapor deposition of two materials, which have been used for preparation of CdSe nanocluster layers, has been also applied for nonepitaxial growth of nano-sized metallic clusters embedded in insulating matrix. Several assumptions have been previously mentioned for the mechanism of the metal nanoparticle formation. It has been assumed [186] that the existing valleys

of the thin film a-Si:H surface support clustering of Mo and W. Also, the cluster formation at most of the transition and noble metals on insulator has been attributed [196] to the large difference of surface energies. The partial wetting character of Ga with respect to SiO_x has been considered [190] as responsible for the formation of liquid Ga nanoclusters on the "smooth" surface of SiO_x ultra-thin films. The cross-sectional electron micrograph of SiO_x/CdSe multilayers have shown [95, 96] that when in the SiO_x/Ga MLs the ultra-thin metal layers are replaced by CdSe semiconductor layers, the latter is continuous rather than "island" type. However, when the SiO_x (GeS_2, ZnSe) surface is "rough," CdSe particle formation takes place, as the spatial distribution of nanoparticles follows the surface morphology of the SiO_x (GeS_2) films. This implies that different mechanisms should be responsible for the metal and semiconductor nanoparticle formation. Obviously, the assumption for difference of surface energies cannot explain deposition of continuous CdSe layers on a relatively "smooth" surface and discontinuous ones on a relatively "rough" surface. In addition, one could not expect the same partial wetting character of CdSe with respect to SiO_x, GeS_2, and ZnSe.

The HRTEM and X-ray diffraction results showed that the surface roughness of SiO_x layers in the SiO_x/CdSe MLs having layer thicknesses <10 nm is relatively small. Therefore, it has been assumed [95] that a homogeneous embryo formation takes place at the first stage of CdSe deposition on the "smooth" SiO_x surface. A gradual increase of the nanoparticle size and coalescence as well as new embryo formation can be expected further; CdSe layers become continuous at a nominal thickness ≥2.5 nm. The observed spatial distribution of the CdSe NCs in the composite films indicates that surface roughness plays an important role in the CdSe nanoparticle formation. It has been suggested [199] that at the very beginning of CdSe deposition on a "rough" oxide or chalcogenide surface embryos were formed at those surface positions at which the curvature and lattice stress are the greatest. The relatively narrow size distribution implies that further CdSe deposition does not create new embryos but leads mainly to an increase of the nanoparticle size. As the positions with great surface curvature and lattice stress are disposed accidentally, some embryos are created rather closely and, for this reason, some nanoclusters are in contact.

In the end of this section, we would like to pay attention to several slightly different applications of the multilayer approach. The first one is a size-controlled nc-Si synthesis, which has been recently realized in SiO/SiO_2 MLs [212, 213]. Thermally induced phase separation has been applied in order to grow Si NCs in the SiO layers. By this technique, a separate control of crystallite size and density is possible. This approach allows fabrication of Si nanocrystals of a well-defined size and rather high density, which is very important for fabrication of light-emitting devices. Germanium island (6–7 nm in diameter) formation with a high density (2×10^{12} cm^{-2}) has been reported [214] on ultra-thin SiO_2 films. It is important that each dot in this multilayer structure had a boundary with the SiO_2 film and, also, with a Si spacer layer.

Another application of the multilayer approach is in preparation of metal-doped films using metal dissolution in ammealed MLs [215]. Multilayers have been prepared by alternative metal (Ag) and chalcogenide ($As_{33}S_{67}$) vapor deposition in vacuum. Such MLs have shown a greater sensibility to illumination and a larger photo-dissolution rate in comparison to the conventional double-layer structure $Ag/As_{33}S_{67}$. The ML structure has been used to fabricate a phase grating for operation in the infrared region. Furthermore, Mo/Si MLs deposited using electron beam evaporation were used to study the formation of silicides at high temperatures [216]. It has also been demonstrated [217] that the introduction of SiO_2 layers in the interface of Mo/Si multilayers was quite effective in improving the heat stability of the MLs. An asymmetric SiO_2 layer thickness of 0.5 and 1.5 nm has been found to be most favorable at the Si-on-Mo interface an Mo-on-Si interface, respectively.

All multilayers described in Section 4, which contain discontinuous metal or semiconductor nanoparticle layers as well as techniques used for their preparation, are listed in Table 3.

Table 3. Multilayers with discontinuous metal or semiconductor layers.

Discontinuous layer/matrix	Preparation technique
Co/Al_2O_3	radio frequency sputtering [193–195]
Ga/SiO_x	thermal evaporation [190]
Mo/a-Si:H	radio frequency sputtering [186]
Ti/a-Si:H	radio frequency sputtering [186]
V/a-Si:H	thermal evaporation of V, glow discharge a-Si:H [186]
Ti/a-Si	[60]
CdSe/a-SiO_x	thermal evaporation [97, 98, 188, 199, 201]
CdSe/a-GeS_2	thermal evaporation [97, 98, 200, 201]
CdSe/ZnSe	thermal evaporation [97, 98, 201]
a-nanoparticle-Si/SiO_2	hot-wire CVD [202]
a-nanoparticle-Si/Si_3N_4	hot-wire CVD [202]
a-Si/Al_2O_3	laser beam evaporation [203]
SiO/SiO_2	thermal evaporation, high temperature annealing [213, 214]

5. CONCLUDING REMARKS

The results described in this article show that preparation of high-quality amorphous multilayers and superlattices and proper post-treatment can be applied for fabrication of semiconductor and metal nanoparticle layers buried in various thin film matrices. The production of amorphous multilayers is realized mainly by using vapor deposition techniques that are widely available, inexpensive, and do not require rather long times for sample deposition. In addition, the application of the step-by-step deposition approach ensures good periodicity and abrupt interfaces of the multilayers. Vapor deposition techniques have the following advantages:

1. The size of nanoparticles can be effectively controlled by predesigned layer thickness during sample deposition.
2. Surface contamination of nanoparticles, which significantly affects their electronic structure, photoluminescence, etc., is avoided.
3. The whole ML structure has a high stability.

However, using the described multilayer approach, it should be kept in mind that while at semiconductor-oxide (nitride, organic) interfaces, material intermixing is negligible, it might be rather strong at semiconductor-semiconductor interfaces. Such intermixing may affect many properties of materials comprising semiconductor-semiconductor interfaces and should be taken into account in all research and application studies.

The success in fabrication of continuous and discontinuous nanoparticle layers in multilayer structure makes possible fundamental investigations in the field of the classic and quantum physics. A great number of new phenomena have been observed that raise new questions for discussion [32, 96, 218]. The observed new size-related properties of these materials allow a number of applications in opto-electronics (light-emitting devices, color filters, photoresistors, etc.), nonlinear optics (transducers, optical keys, waveguides, etc.), and solar energy conversion. The possibilities for xerographic applications and information processing and storage are also subjects of investigations.

GLOSSARY

Continuous nanoparticle layer A layer consisting of densely packed nanoparticles, which are in close contact.

Discontinuous (island type) nanoparticle layer A layer consisting of nanoparticles, a part of which are isolated while others may be in contact with each other.

Multilayer A film consisting of three or more thin layers.

Nanoparticle Crystalline or amorphous particle of a given material and varying shape having a nanometer size in all dimensions.

Periodic (regular) ML A ML built of two kinds of alternating amorphous or/and crystalline thin layers. Each kind of layers is characterized by constant thickness.

Superlattice A periodic ML having ultra thin parallel layers and smooth interfaces.

ACKNOWLEDGMENTS

The author thanks all colleagues from her laboratory as well as Associate Professor C. Raptis from the National Technical University of Athens, Greece, and Dr. H. Hofmeister from the Max-Plank Institute, Halle, Germany, who work with the author on preparation and/or investigation of multilayer structures including nanoparticle semiconductor layers. Dr. Nesheva is also very grateful to the Royal Society for granting collaborative studies in the field with Dr. C. Main and Dr. S. Reynolds from the University Abertay Dundee, Dundee, United Kingdom.

REFERENCES

1. R. Dingle, *Festkoeprobleme* 15, 21 (1975).
2. B. A. Tavger and V. Ya. *Demishovskii, Usp.Fiz.Nauk* 96, 61 (1968); *Sov. Phys. - Usp.* (English transl.) 11, 644 (1969).
3. A. Ya. Shik, *Fiz. Tekh. Poluprovodn.* 8, 1841 (1974); *Sov. Phys. - Semicond.* (English transl.) 8, 1195 (1975).
4. L. Esaki and R. Tsu, *IBM J. Res. Develop.* 14, 61 (1970).
5. L. Esaki, in "Recent Topics in Semiconductor Physics" (H. Kamimura and Y. Toyozawa, Eds.), World Scientific, Singapore, 1983.
6. A. G. Milnes, *Solid State Electron.* 29, 99 (1986).
7. S. Luryi and A. Kastalsky, *Appl. Phys. Lett.* 45, 164 (1984).
8. F. Capasso, *J. Vac. Sci. Technol. B* 1, 457 (1983).
9. D. A. B. Miller, D. S. Chemla, T. O. Damen, A. C. Gossard, W. Wiegmann, I. H. Wood, and C. A. Burrus, *Appl. Phys. Lett.* 45, 13 (1984).
10. E. J. Caine, S. Subbana, H. Kroemer, J. Mers, and A. Y. Cho, *Appl. Phys. Lett.* 45, 1123 (1984).
11. C. Weisbuch and B. Vinter, "Quantum Semiconductor Structures," Academic Press, San Diego, 1991.
12. H. Sakaki, *Solid State Commun.* 92, 119 (1994).
13. B. Abeles and T. Tiedje, *Phys. Rev. Lett.* 51, 2003 (1983).
14. T. M. John, P. Veit, R. Anton, and T. Drüsedau, *Thin Solid Films* 296, 69 (1997).
15. N. Maley and J. S. Lannin, *Phys. Rev. B* 31, 5578 (1985).
16. S. Miyazaki and M. Hirose, *Phil. Mag. B* 60, 23 (1989).
17. T. Kuwano, H. Tarui, T. Takahama, M. Mishikuni, Y. Hishikawa, N. Nakamura, S. Tsuda, S. Nakano, and M. Ohnishi, *J. Non-Cryst. Solids* 97–98, 289 (1987).
18. N. Inoe, Y. Kwamura, and Karimoto, in "Handbook on Nanophase Materials." (A. N. Goldstein, Ed.), Marcel Dekker, New York, 1997, p. 83.
19. W. T. Tsang, *IEEE J. Quant. Electron.* QE-20 1119 (1984).
20. E. Vateva, in "Physics and Applications of Non-Crystalline Semiconductors in Optoelectronics," NATO ASI Series, (A. Andriesh and M. Bertolotti, Eds.), Kluwer Academic Publishers, London, 1997, p. 61.
21. J. Takada and H. Fritzsche, *J. Non-Cryst. Solids* 97–98, 289 (1987).
22. K. Mui and F. W. Smith, *J. Non-Cryst. Solids* 97–98, 975 (1987).
23. R. Ionov and D. Nesheva, *Solid State Commun.* 82, 959 (1992).
24. S. A. Koehler, *Phil. Mag. B* 77, 27 (1998).
25. F. Wang, T. Fischer, T. Muschik, and R. Schwarz, *Phil. Mag. B* 68, 737 (1993).
26. D. Nesheva, D. Arsova, and Z. Levi, *Phil. Mag. B* 70, 205 (1994).
27. R. Ionov and D. Nesheva, *Superlattices & Microstructures* 11, 439 (1992).
28. D. Nesheva, D. Arsova, E. Skordeva, "Proc. 7th International School on the Condensed Matter Physics, Varna 92" (J. Marshall, N. Kirov, and A. Vavrek, Eds.), World Science Publishing, Singapore, 1993, p. 394.
29. I. Pereyra, M. N. P. Carreno, R. K. Onmori, C. A. Sassaki, A. M. Andrade, and F. Alvarez, *J. Non-Cryst. Solids* 97–98, 871 (1987).
30. R. I. Ionov, *Solid State Commun.* 85, 869 (1993).
31. H. Hamanaka, S. Konagai, K. Murayama, M. Yamaguchi, and K. Morigaki, *J. Non-Cryst. Solids* 198–200, 808 (1996).
32. A. D. Yoffe, *Adv. Phys.* 42, 173 (1993).
33. K. Lu, *Matt. Sci. Eng.* R 16, 161 (1996).
34. S. Veprek, *Thin Solid Films* 297, 145 (1997).
35. A. D. Yoffe, *Adv. Phys.* 50, 1 (2001).
36. K. L. Choy, in "Handbook on Nanostructured Materials and Nanotechnology" (H. S. Nalwa, Ed.), Academic Press, San Diego, 2000, Vol. 3, p. 534.
37. K. Reichelt and X. Jiang, *Thin Solid Films* 191, 91 (1990).
38. T. C. Tisone and P. D. Cruzan, *J. Vac. Sci. Technol.* 12, 677 (1975).
39. T. Suntola, in "Handbook on Thin Film Processes Technology" (D. A. Glocker and S. I. Shah, Eds.), Institute of Physics Publishing, Bristol, 1995, p. B1:5.1.
40. K. Reichelt, *Vacuum* 38, 1083 (1988).
41. S. A. Barnett and I. T. Ferguson, in "Handbook on Thin Film Processes Technology" (D. A. Glocker and S. I. Shah, Eds.), Institute of Physics Publishing, Bristol, 1995, p. A2.
42. J. R. Arthur, *J. Appl. Phys.* 39, 4032 (1968).
43. Y. J. Cho and K. Y. Cheng, *Appl. Phys. Lett.* 38, 360 (1981).

44. M. Rabe, M. Lowisch, F. Kreller, and F. Henneberger, *Phys. Status Solidi* (*B*) 202, 817 (1997).
45. Z. H. Lu, D. J. Lockwood, and J.-M. Baribeau, *Nature* 378, 258 (1995).
46. D. J. Lockwood, Z. H. Lu, and J.-M. Baribeau, *Phys. Rev. Lett.* 76, 539 (1996).
47. B. T. Sullivan, D. J. Lockwood, H. J. Labbe, and Z. H. Lu, *Appl. Phys. Lett.* 69, 3149 (1996).
48. S. V. Novikov, J. Sinkkonen, O. Kilpela, and S. V. Gastev, *J. Cryst. Growth* 15, 1471 (1997).
49. S. V. Novikov, J. Sinkkonen, O. Kilpela, and S. V. Gastev, *J. Vac. Sci. Technol. B* 175–176, 514 (1997).
50. L. Khriachtchev, M. Rasanen, S. Novikov, O. Kilpela, and J. Sinkkonen, *J. Appl. Phys.* 86, 5601 (1999).
51. L. Khriachtchev, S. V. Novikov, and J. Lahtinen, *J. Appl. Phys.* 92, 5856 (2002).
52. Y. Kanemitsu and T. Kushida, *Appl. Phys. Lett.* 77, 3550 (2000).
53. S. Nihohyanagi, K. Nishimoto, and Y. Kanemitsu, *J. Non-Cryst. Solids* 299–302, 1095 (2002).
54. M. Zacharias, J. Blaesing, K. Hirschman, L. Tsybeskov, and M. Fauchet, *J. Non-Cryst. Solids* 299–302, 640 (2002).
55. M. Schmidt, J. Heitmann, R. Scholz, and M Zacharias, *J. Non-Cryst. Solids* 299–302, 678 (2002).
56. Zh. Ma, L. Wang, K. Chen, W. Li, L. Zhang, Y. Bao, X. Wang, J. Xu, X. Huang, and D. Feng, *J. Non-Cryst. Solids* 299–302, 648 (2002).
57. Y. H. Kwon, C. J. Park, W. C. Lee, D. J. Fu, Y. Shon, T. W. Kand, C. Y. Hong, and H. Y. Cho, *Appl. Phys. Lett.* 80, 2502 (2002).
58. A. Morimoto, K. Mizushima, and T. Shimizu, *J. Non-Cryst. Solids* 97–98, 943 (1987).
59. A. Asana, T. Ichimura, M. Ohsawa, H. Sakai, and Y. Uchida, *J. Non-Cryst. Solids* 97–98, 971 (1987).
60. R. H. Willens, *Appl. Phys. Lett.* 49, 63 (1986).
61. X. Wu, A. M. Bittner, K. Kern, Ch. Eggs, and S. Veprek, *Appl. Phys. Lett.* 77, 645 (2000).
62. E. L. Zeballos-Velasquez and M. C. A. Fantini, *J. Non-Cryst. Solids* 209, 175 (1997).
63. L. Tsybeskov, K. D. Hirschman, S. P. Duttagupta, P. M. Fauchet, M. Zacharias, J. P. McCaffrey, and D. J. Lockwood, *Phys. Status Solidi* (*A*) 165, 69 (1998).
64. V. Vinciguerra, G. Franzo, F. Priolo, F. Icona, and C. Spinella, *J. Appl. Phys.* 87, 8165 (2000).
65. R. Marzurczyk, M. Gazicki, and T. Wagner, *Physica E* 15, 65 (2002).
66. I. Honma, H. Hotta, K. Kawai, H. Komiyama, and K. Tanaka, *J. Non-Cryst. Solids* 97–98, 947 (1987).
67. S. M. Prokes, W. E. Carlos, S. Veprek, and Ch. Ossadnik, *Phys. Rev. B* 58, 15632 (1998).
68. M. Zacharias, F. Stolze, T. Druesedau, and W. Bock, *Phys. Status Solidi* (*B*) 189, 409 (1995).
69. H. Freistedt, F. Stolze, M. Zacharias, J. Blaesing, and T. Druesedau, *Phys. Status Solidi* (*B*) 193, 375 (1996).
70. Y. Chen, G. Z. Ran, L. Dai, B. R. Zhang, G. G. Qin, Z. C. Ma, and W. H. Zong, *Appl. Phys. Lett.* 80, 2496 (2002).
71. B. Averboukh, R. Huber, K. W. Cheah, Y. R. Shen, G. G. Qin, Z. C. Ma, and W. H. Zong, *J. Appl. Phys.* 92, 3564 (2002).
72. E. Maruyama, *Jpn. J. Appl. Phys.* 21, 213 (1982).
73. T. Ogino and Y. Mizushima, *Jpn. J. Appl. Phys.* 22, 1647 (1983).
74. H. J. Trodahl, M. W. Wright, and A. Bittar, *Solid State Commun.* 59, 699 (1986).
75. H. Shirai, Sh. Oda, T. Nakamura, and I. Shimizu, *Jpn. J. Appl. Phys.* 26, 991 (1987).
76. R. Ionov, D. Nesheva, and D. Arsova, *J. Non-Cryst. Solids* 137–138, 1151 (1991).
77. E. Vateva, R. Ionov, D. Nesheva, and D. Arsova, *Phys. Status Solidi* (*A*) 128, K23 (1991).
78. R. Ionov and D. Nesheva, *Thin Solid Films* 213, 230 (1992).
79. A. Shterr and A. Kikineshy, *Ukrainian J. Phys.* 35, 599 (1990) (in Russian).
80. A. Imre, V. Fedor, M. Kis-Varga, A. Mishak, and M. Shipljak, *Vacuum* 50, 507 (1998).
81. A. Kikineshy, V. Palyok, and M. Shipljak, *Nanostructured Materials* 12, 417 (1999).
82. E. Vateva and I. Georgieva, *J. Non-Cryst. Solids* 114, 124 (1989).
83. E. Vateva and I. Georgieva, *J. Non-Cryst. Solids* 164–166, 865 (1993).
84. E. Vateva and I. Georgieva, in "Physical Phenomena in Non-Crystalline Semiconductors" (V. V. Himinetz and N. I. Dovgoshey, Eds.), UGU, Uzhgorod, 1989, v. 2, p. 22.
85. E. Vateva and G. Tschaushev, *J. Optoelectronics Advanced Materials* 1, (1999).
86. K. Tanaka, Y. Ichimura, and M. Komasaki, *Thin Solid Films* 189, 51 (1990).
87. Z. Zhang, R. Cheng, and H. Fritzsche, *J. Non-Cryst. Solids* 97–98, 923 (1987).
88. E. Vateva and D. Nesheva, *J. Non-Cryst. Solids* 191, 205 (1995).
89. P. Photopoulos, A. G. Nassiopolou, D. N. Kouvatsos, and A. Travlos, *Appl. Phys. Lett.* 76, 3588 (2000).
90. P. Photopoulos and A. G. Nassiopolou, *Appl. Phys. Lett.* 77, 1816 (2000).
91. B. V. Kamenev, *J. Appl. Phys.* 90, 5735 (2001).
92. D. Nesheva, C. Raptis, and Z. Levi, *Phys. Rev. B* 58, 7913 (1998).
93. M. Popescu, F. Sava, A. Lorinczi, E. Vateva, D. Nesheva, G. Tchaushev, I. N. Mihailesku, P.-J. Koch, S. Obst, and H. Bradaczek, *Proc. SPIE* 3405, 964 (1998).
94. D. Nesheva, Z. Levi, and V. Pamuckchieva, *J. Phys.: Condens. Matter.* 12, 3967 (2000).
95. D. Nesheva, *J. Optoelectronics and Advanced Materials* 1, 13 (1999).
96. D. Nesheva, in "Handbook of Surfaces and Interfaces of Materials" (H. S. Nalwa, Ed.), Academic Press, San Diego, 2001, Vol. 3, Chapter 6, p. 239.
97. D. Nesheva, H. Hofmeister, Z. Levi, and Z. Aneva, *Vacuum* 65, 109 (2002).
98. D. Nesheva, Z. Levi, I. Bineva, and H. Hofmeister, in "Nanostructured Materials: Selected Synthesis Methods, Properties and Applications" (P. Knauth and J. Schoonman, Eds.), Kluwer Academic Publishers, Boston, 2002.
99. D. Nesheva, E. Vateva, Z. Levi, and D. Arsova, *Phil. Mag. B* 72, 67 (1995).
100. D. Nesheva, I. P. Kotsalas, C. Raptis, and E. Vateva, *J. Non-Cryst. Solids* 224, 283 (1998).
101. I. Honma, H. Komiyama, and K. Tanaka, *Phil. Mag.* 60, 3 (1989).
102. P. V. Santos, M. Hundhausen, L. Ley, and C. Viczian, *J. Appl. Phys.* 69, 778 (1991).
103. P. Wickboldt, D. Pang, J. H. Chen, H. M. Cheong, and W. Paul, *J. Non-Cryst. Solids* 198–200, 813 (1996).
104. H. Xu, Y. Wang, and G. Chen, *Phys. Stat. Solidi* (*A*) 143, K87 (1994).
105. M. Popescu, F. Sava, A. Lorinczi, E. Vateva, D. Nesheva, P.-J. Koch, T. Gutberlet, W. Uebach, and H. Bradaczeck, *Solid State Commun.* 103, 431 (1997).
106. A. Kikineshy, *Optical Engineering* 34, 1040 (1995).
107. A. Kikineshy and A. Mishak, in "Physics and Applications of Non-Crystalline Semiconductors in Optoelectronics," NATO ASI Series (A. Andriesh and M. Bertolotti, Eds.), Kluwer Academic Publishers, London, 1997, p. 249.
108. M. Marjan, A. Kikineshy, and A. Mishak, *Phil. Mag. B* 68, 689 (1993).
109. D. Nesheva and D. Arsova, *Phys. Status. Solidi* (*A*) 176, R3 (1999).
110. D. Nesheva, C. Raptis, and D. Arsova, *J. Appl. Phys.* 86, 4964 (1999).
111. E. L. Zeballos-Velasquez and M. C. A. Fantini, *J. Non-Cryst. Solids* 209, 175 (1997).

112. J. Kolodzey, P. Hanesch, T. Fisher, R. Schwarz, G. Zorn, and H. Goebel, *Mat. Sci. Eng. B* 11, 43 (1992).
113. L. Tsybeskov, K. D. Hirschman, S. P. Duttagupta, P. M. Fauchet, M. Zacharias, J. P. McCaffrey, and D. L. Lockwood, *Apl. Phys. Lett.* 72, 43 (1998).
114. K. Tanaka, I. Honma, H. Tamaoki, and H. Komiyama, *MRS Proc.* 118, 343 (1988).
115. S. Miyazaki, Y. Ihara, and M. Hirose, *J. Non-Cryst. Solids* 97–98, 887 (1987).
116. P. D. Persans, A. Ruppert, and H. J. Trodahl, *J. Non-Cryst. Solids* 102, 130 (1988).
117. M. Zacharias, J. Blaesing, P. Veit, L. Tsybeskov, K. D. Hirschman, and P. M. Fauchet, *Appl. Phys. Lett.* 74, 2614 (1999).
118. M. Zacharias, J. Blaesing, P. Veit, K. D. Hirschman, L. Tsybeskov, and P. M. Fauchet, in "Quantum Confinement: Nanostructures" (D. Lockwood, Ed.), The Electrochenical Society, Pennington, NJ, 1998, PV 98-19.
119. D. Nesheva, Z. Levi, Z. Aneva, V. Nikolova, and H. Hofmeister, *J. Phys.: Condens. Matter* 12, 751 (2000).
120. J. Gonzalez-Hernandez and R. Tsu, *Appl. Phys. Lett.* 42, 90 (1983).
121. X. Huang, Zh. Li, W. Wu, K. Chen, X. Chen, and Zh. Liu, *J. Non-Cryst. Solids* 198–200, 821 (1996).
122. M. Zacharias, L. Tsybeskov, K. D. Hirschman, P. M. Fauchet, J. Blaesing, P. Kohlert, and P. Veit, *J. Non-Cryst. Solids* 227–230, 1132 (1998).
123. X. L. Wu, G. G. Siu, S. Tong, F. Yan, S. S. Jiang, X. K. Zhang, and D. Feng, *Appl. Phys. Lett.* 69, 523 (1996).
124. G. V. M. Willams, A. Bittar, and H. J. Trodahl, *J. Appl. Phys.* 67, 1874 (1990).
125. M. Zacharias and P. Streitenberger, *Phys. Rev. B* 62, 8391 (2000).
126. K. J. Chen, X. F. Huang, J. Xu, and D. Feng, *Appl. Phys. Lett.* 61, 2069 (1992).
127. K. J. Chen, J. G. Jiang, X. F. Huang, Z. F. Li, and X. X. Qu, *J. Non-Cryst. Solids* 164–166, 853 (1993).
128. K. J. Chen, X. F. Huang, J. Xu, and D. Feng, *Appl. Phys. Lett.* 66, 722 (1998).
129. M. Wang, X. Huang, J. X. Xu, W. Li, Zh. Liu, and K. Chen, *Appl. Phys. Lett.* 72, 249 (1995).
130. W. Wu, X. F. Huang, K. J. Chen, J. B. Xu, X. Gao, J. Xu, and W. Li, *J. Vac. Sci. Technol. A* 17, 159 (1999).
131. D. Bauerle, *Phys. Status Solidi* (*A*) 166, 543 (1998).
132. M. Bianconi, F. J. Fonseca, C. Summone, and G. Fortunato, *J. Non-Cryst. Solids* 137–138, 451 (1991).
133. P. M. Fauchet and I. H. Campbell, *J. Non-Cryst. Solids* 137–138, 271 (1991).
134. A. V. Kolobov, V. Lyubin, T. Yasuda, and K. Tanaka, *Phys. Rev. B* 55, 23 (1997).
135. V. V. Poborchii, A. V. Kolobov, and Ka. Tanaka, *Appl. Phys. Lett.* 72, 1167 (1998).
136. M. Okuda, T. Matsushita, and A. Suzuki, in "Proc. Symp. Phys. Se and Te," Koningstein, Germany, 1979, p. 270.
137. R. Misra, S. K. Tripathi, A. K. Agnihotri, and A. Kumar, *Solid State Sommun.* 77, 797 (1991).
138. S. K. Srivastava, A. K. Aginori, A. Kumar, S. Swarup, and A. N. Nigana, *Jpn. J. Appl. Phys.* 32, 2557 (1993).
139. Li Wang, X. Wang, X. Huang, Zh. Ma, Y. Bao, J. Shi, W. Li, and K. Chen, *J. Non-Cryst. Solids* 299–302, 751 (2002).
140. M. Popescu, F. Sava, A. Lorinczi, P.-J. Koch, T. Gutberlet, W. Uebach, H. Bradaczek, E. Vateva, and D. Nesheva, in "Proc. Int. Semicond. Conf. CAS 96," 19th Ed., Sinaia, Romania, 1996. Vol. 1, p. 301.
141. S. Sen, S. K. Halder, and S. P. S. Gupta, *J. Phys. Soc. Jpn.* 38, 1641 (1975).
142. M. Gusso, L. De Caro, and L. Tapfer, *Solid State Commun.* 101, 665 (1997).
143. H. P. Klug and L. E. Alexander, "X-Ray Diffraction Procedures," Wiley, New York, 1974.
144. X. L. Wu, S. Tong, X. N. Bao, S. S. Jiang, D. Feng, and G. G. Siu, *Appl. Phys. Lett.* 70, 838 (1997).
145. H. Xia, Y. L. He, L. C. Wang, W. Zhang, X. N. Liu, X. K. Zhang, and D. Feng, *J. Appl. Phys.* 78, 6705 (1990).
146. Y. Kanemitsu, T. Ogawa, K. Shirashi, and K. Takeda, *Phys. Rev. B* 48, 4883 (1993).
147. V. P. Afanasjev, A. S. Gudovskikh, J. P. Kleider, K. V. Koughia, A. P. Sazanov, M. E. Gueunier, and E. I. Terukov, *J. Non-Cryst. Solids* 299–302, 1070 (2002).
148. M. Zacharias, R. Weigand, B. Dietrich, F. Stolze, J. Blaesing, P. Veit, T. Druesedau, and J. Christen, *J. Appl. Phys.* 81, 2384 (1997).
149. Y. S. Lim, J. Y. Lee, H. S. Kim, and D. W. Moon, *Appl. Phys. Lett.* 80, 2481 (2002).
150. O. G. Schmidt, U. Denker, S. Christiansen, and F. Ernst, *Appl. Phys. Lett.* 81, 2614 (2002).
151. D. Nesheva, C. Raptis, Z. Levi, I. Bineva, and Z. Aneva, *Asian Journal of Physics* 9, 289 (2000).
152. A. Ekimov, *J. Lumin.* 70, 1 (1996).
153. J. Rodriguez-Viejo, K. F. Jensen, H. Mattoussi, H. J. Michel, B. O. Dabbousi, and M. G. Bawendi, *Appl. Phys. Lett.* 70, 2132 (1997).
154. C. Weisbuch and J. Nagle (Eds.), "Science and Engineering of One and Zero-Dimensional Semiconductors," Plenum, New York, 1990, p. 309.
155. T. Orii, Sh. Kaito, K. Matsuishi, S. Onari, and T. Arai, *J. Phys. Condens. Matter* 9, 4483 (1997).
156. R. K. Jain and R. C. Lind, *J. Opt. Soc. Am.* 73, 647 (1983).
157. P. Roussignol, D. Ricard, J. Lukasik, and C. Flytzanis, *J. Opt. Soc. Am.* 4, 5 (1987).
158. C. Flytzanis, D. Ricard, and M. C. Schanne-Klein, *J. Luminescence* 70, 212 (1996).
159. S. A. Gurevich, A. I. Ekimov, I. A. Kudryavtsev, O. G. Lyubinskaya, A. V. Osinskii, A. S. Usikov, and N. N. Faleev, *Sov. Phys. Semicond.* 28, 830 (1994).
160. A. G. Rolo, O. Conde, and M. J. M. Gomes, *Thin Solid Films* 318, 108 (1998).
161. K. Tsunemoto, H. Nasu, H. Kitayama, A. Kawabuchi, Y. Osaka, and K. Takiyama, *Jpn. J. Appl. Phys.* 28, 1928 (1989).
162. F. Koch and V. Petrova-Koch, *J. Non-Cryst. Solids* 198–200, 840 (1996).
163. M. L. Brongersma, A. Polman, K. S. Min, E. Boer, T. Tambo, and H. A. Atwater, *Appl. Phys. Lett.* 72, 2577 (1998).
164. F. N. Timofeev, A. Aydinli, R. Ellialtioglu, K. Turkoglu, M. Gure, V. N. Mikhailov, and O. A. Lavrova, *Solid State Commun.* 95, 443 (1995).
165. Zh. Ma, X. Liao, J. He, W. Cheng, G. Yue, Y. Wang, and G. Kong, *J. Appl. Phys.* 83, 7934 (1998).
166. F. Iacona, G. Franzo, and C. Spinella, *J. Appl. Phys.* 87, 1295 (2000).
167. K. Murakami, T. Suzuki, T. Makimura, and M. Tamura, *Appl. Phys. A* (Suppl.) 69, S13 (1999).
168. Y. Wakayama, T. Tagami, T. Inokuma, S. Hasegawa, and Sh. Tanaka, *Recent Res. Devel. Crystal Growth Res.* 1, 83 (1999).
169. U. Kahler and H. Hofmeister, *Appl. Phys. Lett.* 75, 641 (1999).
170. U. Kahler and H. Hofmeister, *Optical Materials* 17, 83 (2001).
171. C. Longeaud, J. P. Kleider, P. Roca i Cabarrocas, S. Hamma, R. Meaudre, and M. Meaudre, *J. Non-Cryst. Solids* 227–230, 96 (1998).
172. U. Kreibig and M. Volmer, "Optical Properties of Metal Clusters," Springer, Berlin, 1995.
173. B. Abeles, P. Sheng, M. D. Coutts, and Y. Arie, *Adv. Phys.* 24, 407 (1975).
174. C. Laurent and E. Kay, *Atoms, Molecules, Clusters* 12, 465 (1989).
175. A. Heinrich, J. Schumann, H. Vinzelberg, U. Bruestel, and C. Gladun, *Thin Solid Films* 223, 311 (1993).
176. J. Hajto, A. E. Owen, A. J. Snell, P. G. Le Comber, and M. J. Rose, *Phil. Mag. B* 66, 349 (1991).

177. M. Jafar and D. Haneman, *Phys. Rev. B* 47, 10911 (1993).
178. J. Hajto, B. McAuley, A. J. Snell, and A. E. Owen, *J. Non-Cryst. Solids* 198–200, 825 (1996).
179. T. P. Druesedau and A. N. Panckow, *Phil. Mag. Lett.* 69, 333 (1994).
180. U. Kleineberg, H.-J. Stock, A. Kloidt, B. Schmiedeskamp, U. Heinzmann, S. Hoppe and R. Scholz, *Phys. Status Solidi* (*A*) 45, 539 (1994).
181. Y. Naruse and T. Hatayama, *IEEE Trans. Nucl. Sci.* 36, 1347 (1989).
182. A. N. Panckow, J. Blaesing, and T. P. Druesedau, *J. Non-Cryst. Solids* 164–166, 845 (1993).
183. W. Seifert, N. Carlsson, M. Miller, M. E. Pistol, L. Samuelson, and R. Wallenberg, *J. Progr. Crystal Growth Charact. Mat.* 33, 423 (1997).
184. M. Rabe, M. Lowisch, F. Kreller, and F. Henneberger, *Phys. Status Solidi* (*B*) 202, 817 (1997).
185. E. Kurtz, T. Sekiguchi, Z. Zhu, T. Yao, J. X. Shen, Y. Oka, M. Y. Shen, and Y. Goto, *Superlatt. Microstruct.* 25, 119 (1999).
186. T. P. Drüsedau, A. N. Panckow, and F. Klabunde, *J. Non-Cryst. Solids* 198–200, 829 (1996).
187. B. P. Zhang, T. Yasuda, Y. Segawa, H. Yaguchi, K. Onabe, E. Edamatsu, and T. Itoh, *Appl. Phys. Lett.* 70, 2413 (1997).
188. D. Nesheva and Z. Levi, *Semicon. Sci. Technol.* 12, 1319 (1997).
189. E. Sondergaerd, R. Kofman, P. Cheyssac, and A. Stella, *Surf. Sci.* 364, 467 (1996).
190. D. Tonova, M. Patrini, P. Tognini, A. Stella, P. Cheyssac, and R. Kofman, *J. Phys.: Condens. Matter* 11, 2211 (1999).
191. R. Kofman, P. Cheyssac, A. Aouaj, Y. Lereah, G. Duetscher, T. Ben David, J. M. Pinisson, and B. Bourret, *Surf. Sci.* 303, 321 (1994).
192. Ch. Morawe and H. Zabel, *J. Appl. Phys.* 77, 1969 (1995).
193. F. Fettar, J.-L. Maurice, F. Petroff, L. F. Schelp, A. Vaures, and A. Fert, *Thin Solid Films* 319, 120 (1998).
194. J.-L. Maurice, J. Briatico, J. Carrey, F. Petroff, L. F. Schelp, and A. Vaures, *Phil. Mag. A* 79, 2921 (1999).
195. D. Babonneau, F. Petroff, J.-L. Maurice, F. Fettar, and A. Vaures, *Appl. Phys. Lett.* 76, 2892 (2000).
196. C. T. Campbell, *Surf. Sci. Rep.* 27, 2 (1997).
197. A. Turak, D. Grozea, X. D. Feng, Z. H. Lu, H. Aziz, and A. M. Hor, *Appl. Phys. Lett.* 81, 766 (2002).
198. B. P. Zhang, T. Yasuda, Y. Segawa, H. Yaguchi, K. Onabe, E. Edamatsu, and T. Itoh, *Appl. Phys. Lett.* 70, 2413 (1997).
199. D. Nesheva and H. Hofmeister, *Solid State Commun.* 114, 511 (2000).
200. D. Nesheva, Z. Levi, H. Hofmeistter, Z. Aneva, and G. Bogachev, "Proc. 11th ISCMP, Varna 2000 Materials for Information Technology in the New Millenium" (J. M. Marshall, A. G. Petrov, A. Vavrek, D. Nesheva, D. Dimova-Malinovska, and J. M. Maud, Eds.), Bookcraft, Bath, 2001, p. 436.
201. D. Nesheva, *J. Optoelectronics and Advanced Materials* 3, 885 (2001).
202. K. L. Chopra, "Electrical Phenomena in Thin Films," Mir, Moskwa, 1972.
203. J. Zhu, W. Th. Leach, S. K Stanley, and J. G. Ekerdt, *J. Appl. Phys.* 92, 4695 (2002).
203. P. P. Ong and Y. Zhu, *Physica E* 15, 118 (2002).
204. H. Hofmeister, D. Nesheva, Z. Levi, S. Hopfe, and S. Matthias, "Proc. 11th ISCMP, Varna 2000 Materials for information technology in the new millenium" (J. M. Marshall, A. G. Petrov, A. Vavrek, D. Nesheva, D. Dimova-Malinovska, J. M. Maud, Eds.), Bookcraft, Bath, 2001, p. 432.
205. Y. Golan, L. Margulis, G. Hodes, I. Rubinstein, and J. Hutchison, *Surface Science* 311, L633 (1994).
206. R. Hillebrand, H. Hofmeister, K. Scheerschmidt, and J. Heydenreich, *Ultramicroscopy* 49, 252 (1993).
207. K. Tsunemoto, A. Kawabuchi, H. Kitayama, Y. Osaka, and H. Nasu, *Jpn. J. Appl. Phys.* 29, 223 (1990).
208. Y. Golan, L. Margulis, G. Hodes, I. Rubinstein, and J. Hutchison, *Surface Science* 311, L633 (1994).
209. R. Hillebrand, H. Hofmeister, K. Scheerschmidt, and J. Heydenreich, *Ultramicroscopy* 49, 252 (1993).
210. W. Neumann and H. Hofmeister, "Proc. 11th EUREM," Dublin, 1996, Vol. II, p. 438.
211. J. Tersoff, Yu. Tu, and G. Grinstein, *Appl. Phys. Lett.* 73, 2328 (1998).
212. M. Zacharias, J. Heitmann, R. Scholz, U. Kahler, M. Schmidt, and J. Blaesing, *Appl. Phys. Lett* 80, 661 (2002).
213. J. Heitmann, R. Scholz, M. Schmidt, and M. Zacharias, *J. Non-Cryst. Solids* 299–302, 1075 (2002).
214. A. A. Shklyaev and M. Ichikawa, *Appl. Phys. Lett.* 80 1432 (2002).
215. T. Wagner and P. J. S. Ewen, *J. Non-Cryst. Solids* 266–269, Part 2, 979 (2000).
216. A. K. Srivastava, P. Tripathi, M. Nayak, G. S. Lodha, and R. V. Nandadkar, *J. Appl. Phys.* 92, 5119 (2002).
217. M. Ishino and O. Yoda, *J. Appl. Phys.* 92, 4952 (2002).
218. V. Klimov, in "Handbook on Nanostructured Materials and Nanotechnology" (H. S. Nalwa, Ed.), Academic Press, San Diego, 2000, Vol. 4, p. 451.

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanoparticle Reinforced Thermoplastic Composites

Ming Qiu Zhang, Min Zhi Rong

Zhongshan University, Guangzhou, China

Klaus Friedrich

University of Kaiserslautern, Kaiserslautern, Germany

CONTENTS

1. INTRODUCTION

Polymeric nanocomposites can be considered an important category of organic-inorganic hybrid materials, in which inorganic nanoscale building blocks (e.g., nanoparticles, nanotubes, or nanometer-thick sheets) are dispersed in an organic polymer matrix [1, 2]. They represent the current trend in developing novel nanostructured materials. When compared with conventional composites based on micrometer-sized fillers, the interface between the filler particles and the matrix in a polymer nanocomposite constitutes a much greater area within the bulk material (Fig. 1) and hence influences the composite's properties to a much greater extent, even at a rather low filler loading [3].

From the literature, it is somewhat ambiguous who proposed the term "nanocomposites" for the first time. Different authors gave different opinions [4–6]. Nevertheless, the definition of "nanocomposites" is generally accepted without any argument. That is, it refers to materials consisting of various (two or more) solid phases with different compositions or structures, where at least one dimension is in the nanometer range. It has become familiar worldwide nowadays, especially because of some successful commercial products based on these materials, such as air intake covers and bathroom deodorizers. As far as polymer nanocomposites are concerned, not so many results have been published (Fig. 2), despite a large interest in them. However, polymer nanocomposites have been investigated for a long time, even though the technical term has only recently become popular. For example, extensive studies of carbon black-filled rubbers began before the mid-twentieth century [7]. Semicrystalline polymers can also be considered in some sense another example of nanocomposite materials. They consist of crystalline lamellae (typically 10 nm in width and thickness) dispersed in an amorphous matrix [8].

The recent interest in polymer nanocomposites is based on new achievements expected from their nanostructure, such as unusual mechanical, optical, and magnetic properties. A new degree of freedom for the development of advanced materials with enhanced performance becomes available in this way. In addition, the push for polymer nanocomposites arises from the facts that (1) polymers are still easily processible in many different ways, even with nanoscale solid fillers and (2) nanoparticles are provided by a polymer matrix with a high resistance to any chemical attack. In this way, the unique properties of nanosized objects can be combined with the qualities of a polymer matrix, to fit the requirements of specific applications in diverse areas (Table 1). Examples of polymer nanocomposites include biomimetic ceramic/polymer composites [33, 34], nonlinear optical metal colloid/polymer nanocomposites [35], and intercalated clay/polymer nanocomposites [36, 37].

Numerous procedures for the preparation of polymer nanocomposites have been proposed [38, 39]. They can

*This chapter first appeared in *Handbook of Organic-Inorganic Hybrid Materials and Nanocompsites, Volume* 2: *Nanocomposites*, Edited by H. S. Nalwa. ©2003, American Scientific Publishers.

ISBN: 1-58883-063-2/$35.00

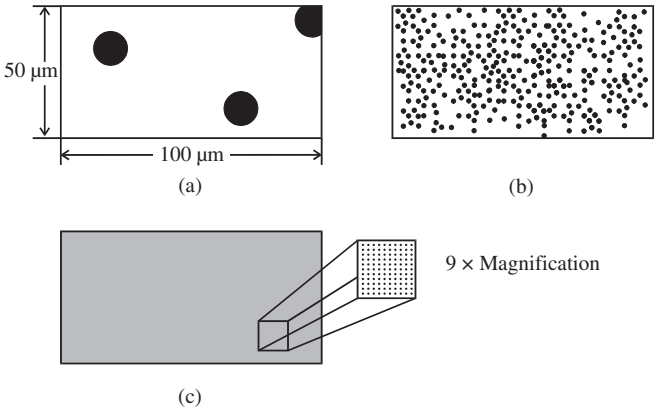


Figure 1. Schematic drawings of the microstructural appearance of typical particulate vs. fine particulate vs. nanoparticulate composites based on electronic microscopic observations. (a) 3 vol.% of particles with 10-μm diameter (2.86 particles within a volume of 50,000 μm^3). (b) 3 vol.% of particles with 1-μm diameter (2860 particles within a volume of 50,000 μm^3). (c) 3 vol.% of particles with 0.1 $\mu m = 100$-nm diameter (2.86 million particles within a volume of 50,000 μm^3). Reprinted with permission from [3], M. Z. Rong et al., *Polymer* 42, 3301 (2001). © 2001, Elsevier Science.

be basically classified in form of the following approaches: (i) direct incorporation of nanoscale building blocks into a polymer melt or solution [22, 40–42]; (ii) *in-situ* generation of nanoscale building blocks in a polymer matrix (e.g., vacuum evaporation of metals, thermal decomposition of precursors, reduction of metal ions through electrochemical procedures, etc.) [15, 43, 44]; (iii) polymerization of monomers in the presence of nanoscale building blocks [45, 46]; and (iv) a combination of polymerization and formation of nanoscale building blocks (e.g., sol–gel method, intercalation of monomers into a layered structure followed by polymerization, etc.) [47–50]. The key issue of these techniques lies in the fact that the geometry, spatial distribution, and volume content of the nanofillers must be effectively controlled through adjustment of the preparation conditions so as to ensure the structural requirements of nano-

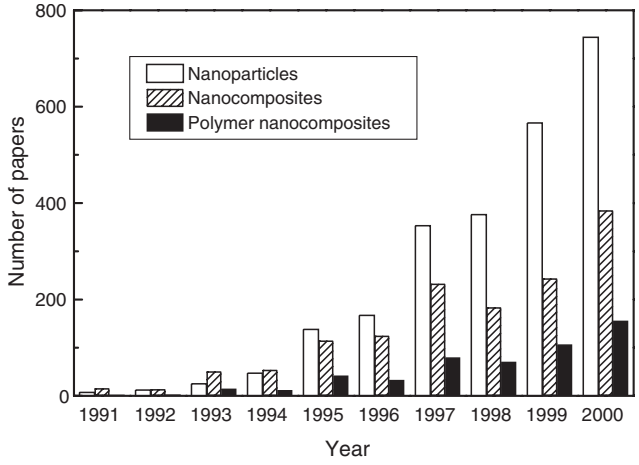


Figure 2. Number of papers published in the period 1991–2000 related to nanostructured materials. Data taken from EI Compendex.

Table 1. Potential applications of polymer nanocomposites based on their specific properties.

Properties	Examples of probable applications	Ref.
Catalytic	Long-life nanocatalysts	[9–12]
Mechanical	Strengthened and toughened polymers	[13, 14]
Magnetic	High-density storage elements, electromagnetic interference shielding materials	[15–18]
Electric	Low-cost conductive coatings, nonlinear resistors, electrostatic discharging filaments, electromagnetic interference shielding materials	[19–21]
Optic	Nonlinear optic materials, optical waveguides, photoelectric materials	[15, 22–26]
Thermal	High thermal conductive materials, low-temperature sintered materials	[19, 27]
Others	Sensors, biomimetic materials, wear-resisting materials, etc.	[28–32]

composites stated above. An ideal nanocomposite should be free of any agglomerates, should contain an optimum filler content, and should maintain at the same time the physicochemical characteristics of its individual components [27]. It is believed that the development of a mixing process for a stable, homogeneous dispersion of the nanophases is crucial for producing a uniform and void-free microstructure in polymer nanocomposites.

It is the aim of this chapter to provide a review of the main results achieved in the preparation and the mechanical behavior of nonlayered, nanoparticle-filled thermoplastic nanocomposites, as manufactured according to the first way (i) cited in the last paragraph. Thermosetting polymer-based composites, like nano-TiO_2/epoxy [51], carbon nanotube/epoxy [52], nano-SiO_2/unsaturated polyester [53], and nano-SiO_2/cycloaliphatic bisepoxide, with the use of an aqueous sodium silicate solution as a starting material [54], are excluded from the main concerns here. Instead, large-scale and low-cost production routes as well as a broad applicability of thermoplastic nanocomposites, intended to be used as structural materials, shall be considered. This can be achieved best by the employment of commercially available, nonlayered nanoparticles and by the use of blending techniques already widely used in the plastics industry.

From an examination of the literature, however, it was surprisingly found that works in this respect are relatively less documented. No global study of such systems has been carried out so far. Most of the papers dealing with thermoplastic systems as candidates for mechanical applications discuss intercalated and exfoliated (i.e., layered) nanocomposites. This phenomenon might be related to the shortcoming of the mixing dispersion method as characterized by the difficulties in (a) breaking up premade nanoparticle agglomerates, and (b) homogeneously arranging individual particles in a polymer matrix. Although great efforts have been made to solve these problems, techniques with sufficiently wide applicability are not yet available. As implied by the statistics in Figure 2, the development of nanocomposites and polymer nanocomposites is growing

moderately, presenting a striking contrast to the case of nanoparticles. This means there is a real need for a revolutionary breakthrough in making and utilizing polymeric nanocomposites.

2. SURFACE MODIFICATION OF NANOPARTICLES

Composite technology is aimed at producing a combination of properties that cannot be achieved with either of the constituents acting alone. Composition, component properties, microstructure, and interactions are the main factors determining the characteristics of composite materials. As compared with the other factors, interactions in composites play the most important role. Two basic types of interactions must be considered: particle–particle and particle–matrix interactions [55]. The effect of the former usually decreases the composite's performance, whereas that of the latter leads to the development of an interphase with properties different from those of the components.

It is known that the interaction between neighboring particles is governed by the competitive contributions made by attraction and repulsion due to van der Waals and electrostatic forces [56]. With decreasing particle size, a reduction in repulsion is much more significant than the decrease in attraction. As a result, attraction between the particles plays the leading role under a certain critical particle diameter. It accounts for the strong tendency of nanoparticles to aggregate.

The structure of a nanoparticle assembly may be classified according to the terminology used in the powder industry. That is, the ultimate primary particles (the smallest particulate pieces) exist in both strongly and weakly bound structures, which are commonly referred to as aggregates and agglomerates, respectively [57]. Agglomerates consist of an assembly of aggregates and have a larger, often more open structure. They behave like a single particle and can be disrupted by considerable force.

In a recent report by Pukanszky and Fekete [58], the forces acting on particulate fillers incorporated in a polymer (fluid or melt) are discussed in detail based on the works of Kendall [59] and Adams and Edmondson [60]. Although many types of forces (e.g., viscous, capillary, electrostatic, etc.) have certain effects, the adhesive force, F_a, which acts as an attractive force, and the hydrodynamic force, taking the role of a separating force, F_h, are the most important ones:

$$F_a = 1.5\pi W_{ab} R \tag{1}$$

$$F_h = 6.12\eta R^2 \dot{\gamma} \tag{2}$$

where R denotes the radius of the particles, W_{ab} is the reversible work of adhesion between the solid and the fluid, η is the viscosity of the fluid, and $\dot{\gamma}$ is the applied shear rate. It was revealed that agglomeration is determined by the relative magnitudes of attractive and separating forces. The predominant factors that influence the homogeneity of composites are the size of the particles, their surface tensions, and the shear forces acting during their homogenization [61]. It should be noted that the interparticle interaction might easily prevail over the shear force and favors the development of agglomerated structures in the case of smaller particles such as nanoparticles. Hence surface modification of the particles becomes critical to minimizing particle–particle interaction and enhancing particle–matrix interaction in composites.

In general, particulate surface modification can be carried out by utilizing physical and chemical interactions between particles and modifiers. The resulting bonds are often a mixture of secondary and chemical bonds. A great understanding of (a) treatment techniques, (b) interfacial characterization, and (c) the effect of interphase on composite properties has been acquired for composites incorporating inorganic fillers [62]. In the following, attention is focused on specific surface modification methods that have been intentionally used to introduce organic coatings onto nanosized filler surfaces.

2.1. Treatment by Physical Interaction

After physical treatment, inorganic fillers can be covered with either low-weight (through a surfactant) or high-weight organic molecules (through an encapsulation with polymer). Since secondary forces (like van der Waals, hydrogen, and electrostatic forces) can theoretically produce an adhesion of 7.0×10^2 to 7.0×10^3 MPa, in comparison with chemical bonds possessing specific adhesive forces of 7.0×10^3 to 7.0×10^4 MPa, the effect of surface modification based on physical interactions is, however, less significant than that of a chemical treatment.

2.1.1. Surfactant

Usually a surfactant contains one or more polar groups and a long aliphatic chain. The principle of surfactant treatment is the preferential adsorption of a polar group of a surfactant to the high-energy surface of a filler by electrostatic interaction. Ionic bonds can also be formed under certain circumstances. A typical example can be cited regarding the surface treatment of $CaCO_3$ with stearic acid [63], which is necessary when producing nano-$CaCO_3$. Stearic acid or its ammonium salt can be added directly during the precipitation step of $CaCO_3$. Electron spectroscopy for chemical analysis (ESCA) studies carried out on the surface of $CaCO_3$ covered with stearic acid showed that ionic bonds have been established between the surfactant molecules and the filler surface (Scheme 1). A further proof of the specific character of the above treatment is supplied by the fact that talc and silica absorb a significantly smaller amount of stearic acid at a unit surface than $CaCO_3$. Evidently, the lack of ionic bonding formation results in a much lower amount of the proportionally bonded molecules. Nevertheless, even for the treated $CaCO_3$, the particles still exist in the form of agglomerates because the interparticulate attraction is far beyond the obstructive effect resulting from the surfactant.

$$CaCO_3 + RCOOH \longrightarrow Ca(OH)(OOCR) + CO_2$$

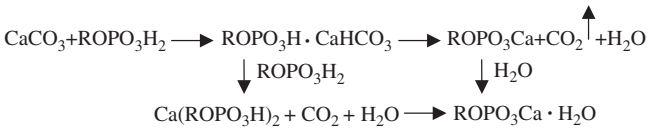
Scheme 1

Papirer et al. related the absorbed amount of stearic acid to the equilibrium concentration of the solution by establishing adsorption isotherms [64]. The quantity of stearic acid

retained by the solid was determined either from the change in concentration of the supernatant solution or directly by analyzing the recovered solid. The physisorption isotherm of stearic acid on $CaCO_3$ (50–100 nm) at 30 °C was found to follow a type II behavior, corresponding to a multilayer adsorption. After extraction of the treated particles with hot toluene for 24 h, the chemisorption isotherm was found to be a Langmuir-type one, which confirms the formation of a monolayer of stearic acid on the surface of $CaCO_3$. The monolayer (100% coverage) suggests a ratio of stearic acid of about 8% by weight. As the surface area of $CaCO_3$, calculated on the assumption of a cross-sectional area of stearic acid of 0.21 nm^2, is 35 m^2/g (in excellent agreement with the value obtained from N_2 adsorption), it was concluded that the stearic acid molecules are fixed perpendicular to the solid surface.

It should be noted that the shield made of C_{18} alkyl chains on the filler surface leads to a reduced surface energy. Contact angle measurements revealed a drastic decrease in surface energy with increasing degree of surface coverage by stearic acid. Before the treatment, $CaCO_3$ possesses a high energy surface. Both the dispersive (γ^d) and the polar (γ^p) components of its surface energy are higher than 50 mJ/m^2. When the particles were covered by a monolayer of stearic acid, the surface energy went down to 22 mJ/m^2, coinciding with the value published for pure stearic acid.

Alkyl dihydrogenphosphate, containing functional groups such as olefine, chloro, methacryloxy, and mercapto, is another kind of effective surface modifier of calcium carbonate fillers [65, 66]. It was shown that alkyl dihydrogenphosphate molecules can react with $CaCO_3$, resulting in a dibasic calcium salt of phosphate. The products are deposited on $CaCO_3$ and provide the surface of $CaCO_3$ with a certain hydrophobicity (Scheme 2), as indicated by its good dispersibility in mineral oil [64]. Relatively larger groups in the phosphates (like octyl dihydrogenphosphate, 3,7-dimethyl-6-octenyl dihydrogenphosphate, etc.) help to enhance the conversion of the agent through a lower hydrophilicity. Furthermore, a treatment with alkyl dihydrogenphosphates always results in an increase in the specific surface area of the treated particles. This can be ascribed to surface erosion and/or calcium salt formation.



Scheme 2

Zhou and co-workers modified nanosized Y-TZP powders (10–20 nm, consisting of 97 mol% ZrO_2 and 3 mol% Y_2O_3) with adipic acid and stearic acid, respectively (Fig. 3) [67]. A covalent bond was established by the chemical reaction between the surface hydroxylic groups and the acid groups of the surfactants. Fourier transform infrared (FTIR) spectroscopy proved that a monolayer of aliphatic chains was coated on the Y-TZP particulates, since the peak representing free carboxyl was no longer detectable. As a result, the flowability of the powders was improved because of the change of their surface polarity.

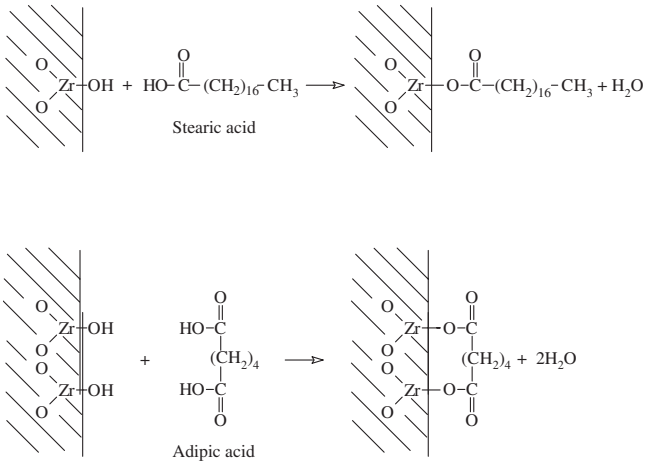
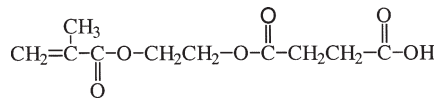
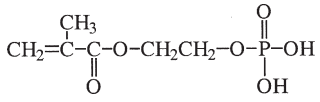
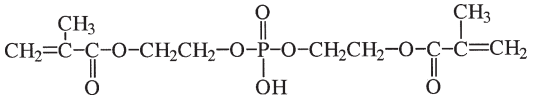


Figure 3. Formation of a monomolecular layer on Y-TZP particles. Reprinted with permission from [67], J. G. Zhou et al., *J. Inorg. Mater.* 11, 237 (1996). © 1996, Shanghai Institute of Ceramics, Chinese Academy of Sciences.

Yoshihara reported the application of hydrophilic dispersants of two-component structures (oxyethylene chains and acidic groups) for the preparation of a stable ultrafine particle dispersion (UFP) (chemical composition unknown, having hydrophilic and cationic nature, 120 nm in size) [68].

The results indicated that the interparticle attraction was reduced because of an adsorbed layer of dispersant on the surface of UFP by a hydrophilic interaction. On the principle of acid–base interaction, the oxyethylene chains might permeate the interface of the UFP agglomerates, and acidic groups contribute to hydrogen bonds and cationic hydroxyl groups on the UFP. Table 2 lists the effect of different dispersants on the average particle size of UFP. Obviously, the effect of the dispersants is not enough to separate the agglomerates completely. Furthermore, it was found that these dispersants could be easily desorbed from the UFP surfaces with the addition of small amounts of H_2O to the dispersion, leading to an increase in the average particle size. That is, reagglomeration occurred with increasing H_2O.

Table 2. Structure of dispersant and average particle size of UFP.

	Reactive dispersant	Particle size[a] (nm)
A:		250
B:		200
C:		195

[a] Dispersal conditions: UFP, 10.0 g; dispersant, 1.0 g; diethylene glycol dimethyl ether, 16.5 g; paint shaker, 3 h.

Source: Reprinted with permission from [68], T. Yoshihara, *Int. J. Adhes. Adhes.* 19, 353 (1999). © 1999, Elsevier Science.

This dispersion system can be illustrated by the correlation among the three components in Figure 4. One of the major problems of the system is that the affinity of the dispersant for the UFP surface is lower than that of H_2O.

To solve the problem, grafting of a polymer to surface hydroxyl groups of the UFP was conducted to inhibit H_2O absorption and reagglomeration of the particles (Fig. 5). Figure 6 demonstrates the concept related to the application of grafting polymer to form hydrophobic layers that have an affinity for the dispersant. It can be expected to obtain a stable dispersion due to an improvement of interaction between the grafted UFP and the dispersant interface. The feasibility of the above proposed approach was subsequently demonstrated by an examination of the stability of the dispersion of poly-(dimethylsiloxane)-grafted UFP in diethylene glycol dimethyl ether. That is, a stable dispersion was kept even after 30 days in the case of polymer-grafted UFP, whereas the untreated UFP was completely precipitated after 1 week.

2.1.2. Encapsulation with Polymer

Encapsulation with preformed polymers or *in-situ* formed polymers can also promote the surface hydrophobicity of inorganic fine particles. In this case, the surface feature of the treated particulates is different from that of surfactant-treated versions because both the surface interaction and the morphology of the encapsulating polymer are more complex.

Hyperdispersant, which has been successfully used to treat inorganic dyestuffs in the paint industry (some of which fall into the category of nanoparticles), is a macromolecular dispersant [69]. Similar to surfactants, polymeric dispersants consist of two major components. One is a functional group, like -OH, $-NH_2$, $-NR^{3+}$, -COOH, $-COO^-$, $-SO_3H$, $-SO^{3-}$, or $-PO_4^{2-}$, which helps to anchor dispersants to the particle surface through hydroxyl and electrostatic bonds. The other is a soluble macromolecular chain, like polyolefine, polyester, polyacrylate, or polyether, which is appropriate for dispersion in different media of low to high polarity. In comparison with traditional surfactants, hyperdispersants have the following advantages: (i) they are anchored to the particles' surface more strongly than are surfactants and hence can hardly be desorbed; (ii) the long polymeric chains

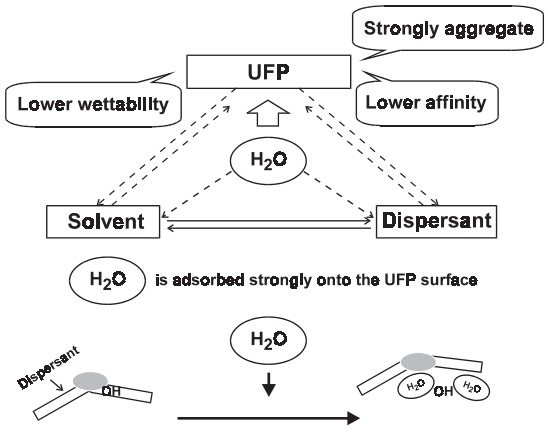


Figure 4. Influence of H_2O on UFP dispersal system. Reprinted with permission from [68], T. Yoshihara, *Int. J. Adhes. Adhes.* 19, 353 (1999). © 1999, Elsevier Science.

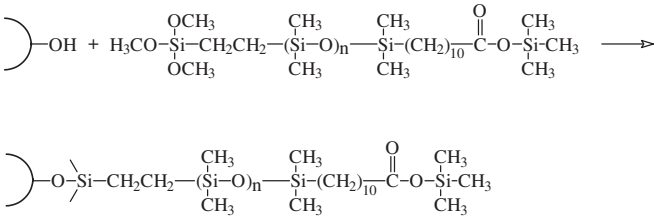


Figure 5. Grafting polymers onto surface hydroxyl groups of UFP. Reprinted with permission from [68], T. Yoshihara, *Int. J. Adhes. Adhes.* 19, 353 (1999). © 1999, Elsevier Science.

can interfere with the reagglomeration of the particles more effectively; (iii) their specific molecular structures, characterized by an A-B or B-A-B block copolymer type, ensure the efficiency of particle isolation and avoid bridging between particles through an identical soluble chain. When nanoparticles are treated, however, it should be noted that hyperdispersants can only encapsulate nanoparticle agglomerates and can hardly penetrate the agglomerates, because of their long macromolecular structure.

In the presence of silica colloid, Yoshinaga and coworkers utilized radical polymerization of vinyl monomers, such as styrene, methyl methacrylate, and 2-hydroxyethyl methacrylate (HEMA), initiated by 2,2′-azobis (2-amidinopropane) dihydrochloride (AAP), to produce SiO_2/polymer composites with a retained particle size of 470 nm [70]. Since the surface of silica colloid particles possesses negative charges on the electric double layer, the initiator AAP can be concentrated, resulting in free radicals on the silica surface. It is thus expected that an efficient polymerization of vinyl monomer would take place.

The binding between the *in-situ* formed polymer and the silica surface is driven by an electrostatic attraction. It was found that the solvent used during polymerization was an important factor influencing the percentage of the attached polymer. At an initial stage, polymerization of vinyl monomers occurred first on the silica surface, leading to a certain degree of hydrophobicity. This simultaneously promoted aggregation of the polymer chains to be formed in the ethanolic bulk solution because of the poor solubility of the polymer in an ethanol-rich solvent. In this context, a solvent with lower miscibility for the polymer should facilitate the attachment of the macromolecular chains to the particles.

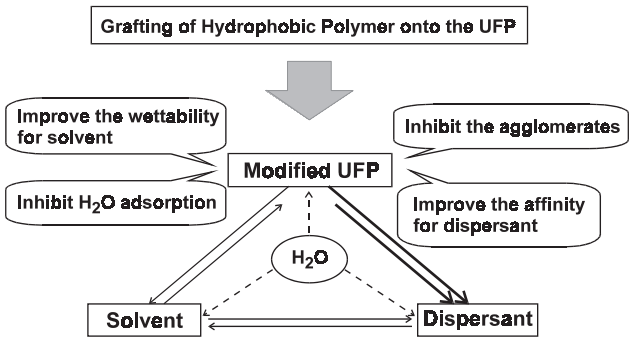


Figure 6. Grafting of hydrophobic polymer. Reprinted with permission from [68], T. Yoshihara, *Int. J. Adhes. Adhes.* 19, 353 (1999). © 1999, Elsevier Science.

Hasegawa and co-workers proposed an encapsulation process capable of uniformly covering fine inorganic powders of submicron size with a filmy polymer [71]. The encapsulation of barium sulfate (0.6 μm) and calcium carbonate (0.4 μm) was attempted by a soapless emulsion polymerization of methyl methacrylate in water (in the presence of the inorganic powders). The encapsulation status of the particles with the polymer formed varied considerably with the species of the initiator and the reaction atmosphere. When potassium persulfate (acting as an initiator) and nitrogen were used, the fine powder surface was partially covered by polymer particles that were the same size as the powder particles. In addition, many polymer chains did not adhere to the powder surface, meaning that the monomers were not used effectively to cover the powders. In the case of a redox initiator under an air atmosphere, however, the powder surface was well encapsulated with a film-like polymer layer. It was believed that the formation of a large number of oligomers with short alkyl chains at an earlier stage of reaction should be responsible for the above difference. On the basis of this mechanism, the authors suggested that the addition of an extremely small quantity of a certain surfactant to the reaction system prior to the polymerization would facilitate the polymerization on the powder surface.

Du et al. [72] demonstrated that when a seed emulsion polymerization was carried out in the presence of nanoparticles, a core-shell structure can be obtained, in which the inorganic particles stay in the middle layer while the polymer is located at inner and outer layers. For the system containing Fe_2O_3 nanoparticles (3–5 nm) and styrene/acrylic acid/butyl acrylate (St/AA/BA) latex particles, for example, the size of the resultant composite microspheres is 80 nm. Changes in the infrared absorption band of carbonyl groups indicated a strong interaction between Fe_2O_3 nanoparticles and the surface groups of the seed latex particles.

Nanoparticles covered by a polymer layer can also be obtained during the manufacture of nanoparticles. For instance, nanocrystalline titania coated with poly (methacrylic acid) (PMAA) was prepared by microwave-induced plasma [73]. Such an *in-situ* treatment is characterized by high effectiveness and homogeneity (Fig. 7). Transmission electron microscopy (TEM) observation showed that the titania particles (rutile in majority with a few anatase) were 10–25 nm in diameter, with a 7-nm-thick PMAA layer.

During the synthesis of metallic and semiconductive particles, polymeric materials are frequently employed as particle stabilizers to prevent agglomeration and growth of the particles. In the present case, however, existing nanoparticle agglomerations must be separated in advance, that is, before encapsulation with a polymer in solution can take place. This means that, under these circumstances, particle deagglomeration, electrostatic stabilization of the solvent, and the interpenetration plus electrostatic stabilization of the polymer play an important role. Gonsalves and Chen illustrated the above points when preparing highly loaded AlN/polyimide nanocomposites (Fig. 8) [74]. In their work, *N*-methylpyrrolidinone (NMP) was used as a solvent, functioning as a medium for the deagglomeration of the nanoparticles and their stabilization through electrostatic interactions. With the addition of poly(amide acid) (PAA), the nanoparticles were further deagglomerated and stabilized through both steric and electrostatic stabilization. Acting as a polyelectrolyte, PAA tends (a) to diffuse into the pores of the secondary structures of the AlN nanoparticles and (b) to be absorbed by the particle surfaces. To prevent any possible destabilization of the homogeneous suspension system, a rapid precipitation method making use of different nonsolvents and suspension/nonsolvent ratios should be used. For the AlN/PAA system, triethylamine was considered the best choice because it completely precipitated out of the composite phase of the suspension.

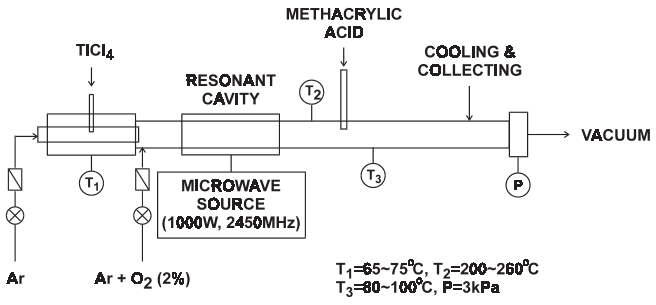


Figure 7. Synthesis of nanocrystalline titania with PMAA coating by microwave plasma. Reprinted with permission from [73], H. B. Liu, *Chemistry* 10, 44 (1997). © 1997, Chinese Society for Chemistry.

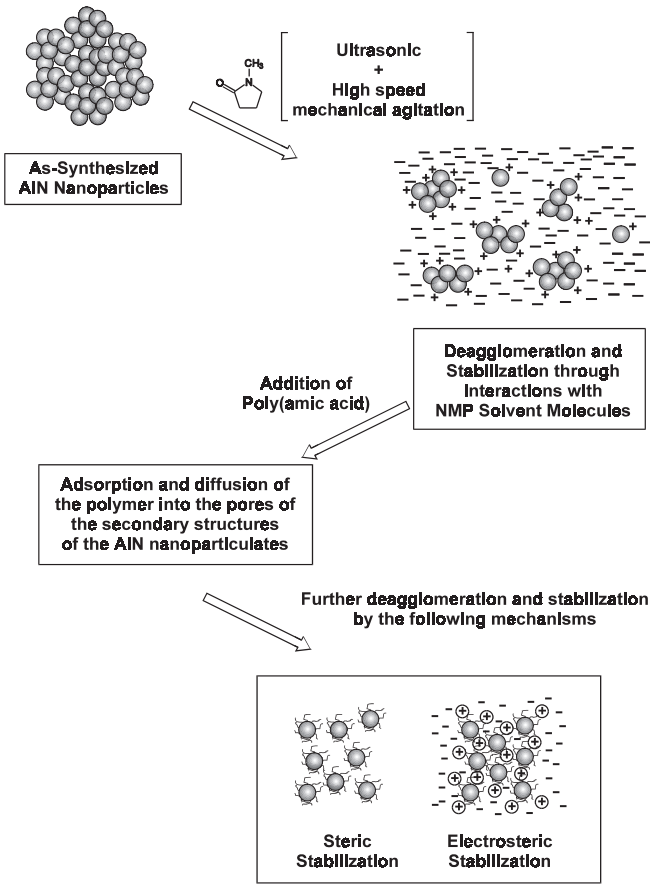


Figure 8. Deagglomeration and stabilization of AlN nanoparticles in a polyelectrolyte solution. Reprinted with permission from [74], K. E. Gonsalves et al., *J. Mater. Res.* 12, 1274 (1997). © 1997, Materials Research Society.

2.2. Treatment by Chemical Reaction

As a surface treatment through chemical reaction can lead to much stronger interactions between modifiers and nanoparticles, much more attention has been paid to this area. In general, a chemical treatment can be divided into two aspects. One deals with the use of lower weight molecules, such as coupling agents, and the other with a grafting of polymer chains. In comparison, chemical techniques are more complicated than physical treatments because physical and chemical absorptions as well as polymerization of the modifiers are involved in the former. As is often the case, a three-dimensional multilayer structure forms on the solid surface, which is somewhat difficult to control and optimize.

2.2.1. Coupling Agent

A variety of coupling agents, such as silane, titanate, zirconate, etc., have been used to improve the adhesion between inorganic fillers and organic matrices. Extensive information on reaction mechanisms and effects of coupling agents is available for microfillers, but less is available for nanoparticles. Actually, there is no evident difference between these two kinds of particles regarding the reaction mechanisms, except that the coupling agents can hardly penetrate nanoparticle agglomerates.

Most commercially available silanes contain three ethoxy or methoxy groups and a linear alkyl chain attached to a single silicon atom. A typical structure of a common silane can be written $RSiX_3$. The functional group X reacts with hydroxyl groups on the filler surface, while the alkyl chain, capped with a reactive group such as amino, vinyl, or methacryloxy, may react with the matrix polymer. Hydroxyls or absorbed water from the inorganic surface can serve to hydrolyze the silane agent. The silanol moieties readily condense to form siloxane bonds that can result in a linear, branched, or cross-linked polysiloxane structure. It is clear that in the case of trifunctional silane, a three-dimensional siloxane multilayer forms on the solid surface, which remains even after extensive washing in boiling water [75]. However, only a portion of X groups in grafted silanes can interact with the particle surface or with other X groups of the reactant because of a relatively low concentration of surface hydroxyls (1–2 OH/nm^2). As a result, a fraction of X groups remains, and these residual groups can influence the properties of both the covering material and the final polymer matrix. The whole process is dependent on the synthesis and treatment conditions.

To illuminate the effect of a silane coupling agent treatment, fumed silica is taken as an example in the following discussion. A structural hierarchy exists for fumed silica, based on four levels [76, 77]. The first level represents the tetrahedrons of SiO_4, which, in turn, form amorphous and almost spherical primary particles of 5–50 nm (second level) containing 10^4 or more SiO_2 units. These particles, linked by -Si-O-Si- bridges (maybe partially) or intermolecular bonds (such as hydrogen bonds), build up relatively stable aggregates (100–500 nm) (third level). Last, the aggregates give rise to loose and relatively unstable agglomerates (>1 μm) (fourth level) due to intermolecular polar and electrostatic bonding effects. The existence of this structural hierarchy of fumed silica can, of course, influence the effect of the surface modifier on the structural features and other characteristics of the silica dispersions in different ways. In other words, after a certain pretreatment, the observed particle size distribution depends on the transformation state between the agglomerates and the aggregates; the latter can be decomposed during pretreatment and subsequently rearranged through different dispersion processes.

Gunko et al. [78] studied the features of fumed silica coverage with silanes from both experimental and theoretical aspects. It was found that a modification with the use of methacryloyloxymethylenemethyl diethoxysilane (MMDES), 3-methacryloyloxypropyltrimethoxysilane (MAPTMS), and 3-aminopropyltriethoxysilane (APTES) strongly changed the ξ(pH) potential. The quantity and the structure of the modifiers controlled the particle size distribution in an aqueous suspension. MMDES gave better results in comparison with its mixture with APTES or with pure MAPTMS, when considering the reduction in particle (agglomerate) size of modified silica at a modifier concentration of over 0.05–0.2 mmol/g. Depending on the pretreatment conditions, grafted silanes can reduce the hydrophilic properties of silica to different extents.

The silylation procedure can be distinguished simply by examining whether the reaction system contains water or an organic agent. In the case of organic silylation, a very high surface concentration of silane molecules would be obtained because the hydroxyl groups on the inorganic particles serve as reactive groups and are involved in grafting organic molecules. However, it should be noted that the chemical reaction of these hydroxyl groups might be selective. Moreover, some chemicals would compete with silane. Baraton and co-workers [79] studied the grafting of hexamethyldisilazane (HMDS) on nanoalumina (100 m^2/g) and nanotitania (50 m^2/g) through *in-situ* FTIR. The results indicated that the reaction of HMDS with different hydroxyl groups on the nanoparticles is selective. Moreover, Si-OH groups can be generated by oxidation of the particles. When acetic acid, which has a strong affinity for hydroxyl groups, was added to these modified surfaces, a partial or total displacement of the grafted HMDS occurred, and indicating selectivity.

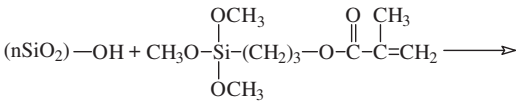
2.2.2. Grafting Polymerization by Chemical Methods

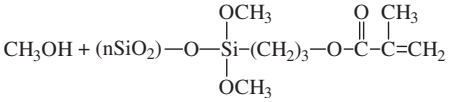
Grafting polymeric molecules to inorganic particle surfaces has some advantages over modification by low molecular surfactants or coupling agents. Polymer-grafted particles can possess desired properties through a proper selection of the species of the grafting monomers and the grafting conditions.

A series of works has appeared in this field of growing interest for purposes of improving the dispersibility of the nanoparticles in solvents and their compatibility in polymer matrices. The grafting polymerization was conducted mostly via two routes: (i) monomers are polymerized from active compounds (initiators or comonomers) covalently attached to the inorganic surface, and (ii) ready-made polymers with reactive end-groups react with the functional groups on the particles. The latter benefits the control of the molecular weight of the grafting polymer. So far, many grafting techniques, such as chemical grafting, irradiation grafting, and

mechanical grafting, have been applied to the modification of nanoparticles. Moreover, various kinds of initiation or polymerization processes, including radical, anionic, and cationic polymerizations, were also undertaken in the grafting investigations.

Grafting from Double Bonds on Particles Espiard and Guyot [80] studied grafting from silica particles with the use of radical processes in an emulsion. The particle surface was first covered with a silane coupling agent containing a polymerizable group (i.e., methacryloxypropyltrimethoxysilane, MPTMS) (Scheme 3). Then the modified silica was engaged in the emulsion polymerization of ethyl acrylate. It was proved that a part of the polymer could be covalently grafted to silica only if the latter has been functionalized. The grafting process took place mostly in the early stage of the polymerization according to the results of the grafting kinetics. The silane-treated silica particles were actually slightly agglomerated when dispersed in water. Their mean diameter amounted to 37 nm, whereas that of the initial particles was only 12–13 nm. The interstitial volume inside these particle agglomerates made it possible to partly locate the polymer there.





Scheme 3

From the surface density of silane molecules on silica and the amount of grafted polymer, the average number of monomer units in a grafted polymer molecule between two grafting points can be estimated, assuming full consumption of the anchored methacrylate. It was found that the estimated value of that number does not exceed 15. This means that the grafted polymer is very close to the particle surface, with rather short loops between the grafting points (Fig. 9).

Oosterling et al. [81] grafted polystyrene (PS) and poly(styrene-block-isoprene) to silica (average particle diameter, 12 nm; specific surface area, 200 m^2/g) by initiating immobilized double bonds on the silica surface with *t*-butylithium and then adding the monomers to the reaction suspension (Fig. 10). The functionalization of silica was done by reacting *p*(*m*)-vinylbenzyltrichlorosilane with the silica surface; however, only 58% of the hydroxyl groups on the silica were consumed (Fig. 11). Therefore, the reaction product had to be treated with HMDS to block the remaining hydroxyl groups, preventing them from interfering with the subsequent anionic polymerization. With the addition of two or more types of monomers one after another, an end-grafted block copolymer was created. The amount of the grafted polymer increases linearly with increasing monomer concentration, as was expected for the anionic polymerization. It appears that the increase in grafting polymer was

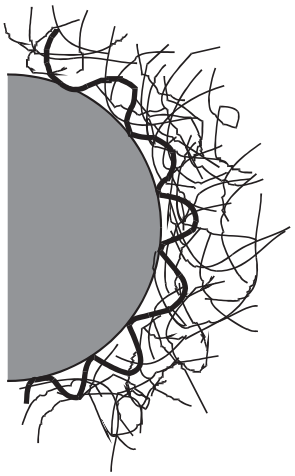


Figure 9. Schematic representation of polymer around a silica particle. Loops of grafted polymer are shown as thick lines attached to the silica, and nongrafted polymer chains are shown as thin lines. Reprinted with permission from [80], Ph. Espiard and A. Guyot, *Polymer* 36, 4391 (1995). © 1995, Elsevier Science.

completely due to an increase in the molecular weight and not to an increase in grafting density.

Isotactic poly(methyl methacrylate) (PMMA) covalently bound to silica (12 nm) can also be synthesized by an initiation of silica-bound methacrylate groups with phenylmagnesium bromide (Fig. 12) [82]. The grafted PMMA appeared to be highly isotactic and had the same molar mass distribution and tacticity as the polymer freely formed in solution. The grafted polymer led to a range of 0.30–0.47 g/g, but only about 0.15 g/g could be obtained if no

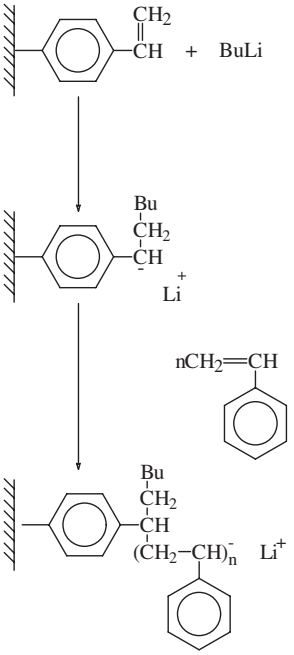


Figure 10. Grafting polymers onto nanoparticles with the use of the double bonds on the particles. Reprinted with permission from [81], M. L. C. M. Oosterling et al., *Polymer* 33, 4394 (1992). © 1992, Elsevier Science.

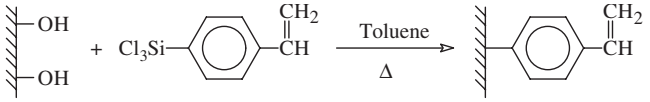


Figure 11. Functionalization of nanosilica with the introduction of double bonds. Reprinted with permission from [81], M. L. C. M. Oosterling et al., *Polymer* 33, 4394 (1992). © 1992, Elsevier Science.

posttreatment with HMDS was applied to the functionalized silica.

Grafting Starting from the Initiator on Particles To obtain polymer-grafted inorganic particles with a higher percentage of grafting, it is preferred to initiate the graft polymerization from the initiating groups introduced on the particles because the grafted polymers can propagate directly from the particles' surface. In a series of articles, Tsubokawa and co-workers [83–88] systematically reported the grafting of various polymers to ultrafine silica and carbon black surfaces by polymerization, starting from the initiating groups previously introduced onto the surfaces. For instance, they have conducted anionic, cationic, and radical graft polymerization of various monomers initiated by potassium carboxylate, acylium or benzylium perchlorate, and peroxyester or azo groups already introduced onto ultrafine silica, respectively. In the following text, methods of introducing initiators onto inorganic particles and the resultant grafting polymerization efficiency will be reviewed briefly.

In consideration of the fact that an anionic ring-opening copolymerization of epoxides with cyclic anhydride can be initiated by potassium carboxylate, Tsubokawa et al. first treated ultrafine inorganic particles (silica (16 nm), titanium oxide (120 nm), and nickel zinc ferrite (15 nm)) with 4-trimethoxysilyl-1,2,5,6-tetrahydrophthalic anhydride (TSPA) to bring acid anhydride groups onto the surface of the particles [83]. Then these acid anhydride groups were neutralized by potassium hydroxide, leading to particles attached with -COOK groups. The results of the grafting of various polymers to the nanoparticles carrying the anionic initiators by the copolymerization of epoxides showed that high percentages of grafting were obtained in all cases.

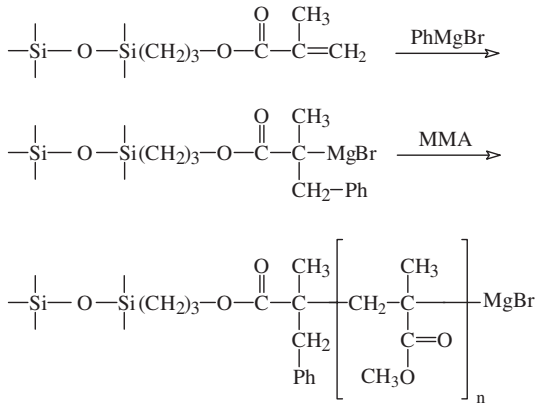


Figure 12. Grafting PMMA onto nanosilica. Reprinted with permission from [82], E. Schomaker et al., *Polym. Commun.* 29, 158 (1988). © 1988, Elsevier Science.

To generate cationic initiators on a silica surface, Tsubokawa et al. also introduced acid anhydride groups onto the particles with the use of TSPA [84]. After that, acyl chloride groups were obtained by treating the surface carboxyl groups with thionyl chloride. Eventually, the introduction of acylium perchlorate (APC) groups onto the silica surface (silica-APC) was achieved by the reaction of acyl chloride groups with silver perchlorate in nitrobenzene at room temperature (Fig. 13). The polymerization of styrene was found to be successfully initiated by silica-APC, and the grafting efficiency reached 64.5%.

As peroxide groups can effectively initiate grafting polymerization of vinyl monomers, Tsubokawa et al. [84–86] have made many attempts to introduce the groups of this kind onto silica (16 nm). In summary, three approaches proved to be feasible. The first one originates from acid anhydride groups that are introduced onto the silica surface by a reaction of silanol groups with TSPA. Then the peroxyester (POE) groups were introduced onto the silica surface by the reaction of surface acid anhydride groups with α, α-dimethyl(4-isopropylbenzyl)hydroperoxide (DIBHP) (Fig. 14). The second is related to the reaction of hydrogen peroxide with chlorosilyl groups, and the latter are introduced by the treatment of silica with thionyl chloride. Both diisopropylbenzene peroxide and *tert*-butyl peroxide groups can be obtained with this method. The third method utilizes a Michael addition reaction of *t*-butylperoxy-2 methacryloyloxyethylcarbonate (HEPO) with amino groups introduced on the silica surface to produce peroxycarbonate groups linked to silica. The experimental data revealed that all of the grafted silica initiated by the above peroxide groups on the treated silica produced a stable colloidal dispersion in organic solvents.

In addition to the peroxide group, azo is another radical initiator commonly in use. Usually it is introduced onto ultrafine particle surfaces in two ways. One deals with the reaction of 4,4′-azobis (4-cyanopentanoic acid) (ACPA) with a surface epoxide group [86], which is introduced by the treatment of the particles with γ-glycidoxypropyltrimethoxysilane (GPS) (Scheme 4). The other is related to the reaction of ACPA with a surface isocyanate group [87], which is introduced by treating the particles with 2,4-tolylene diisocyanate (TDI) (Scheme 5). In addition, by means of

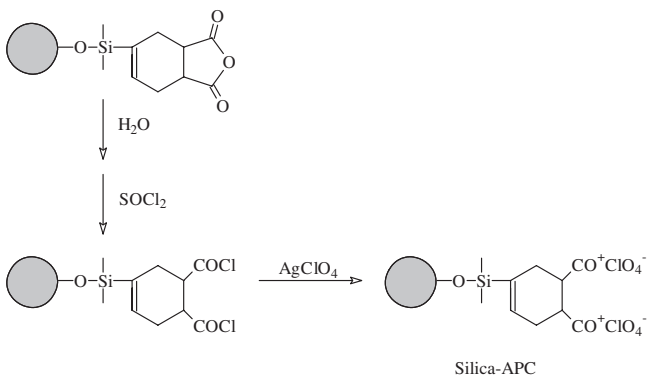


Figure 13. Introduction of acylium perchlorate (APC) groups onto silica. Reprinted with permission from [84], N. Tsubokawa et al., *Polym. Bull.* 31, 456 (1993). © 1993, Springer-Verlag.

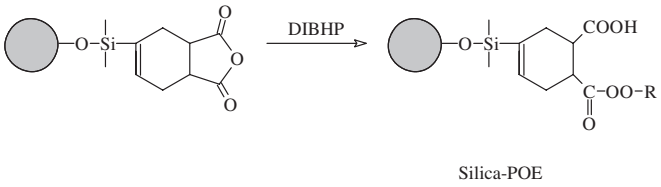


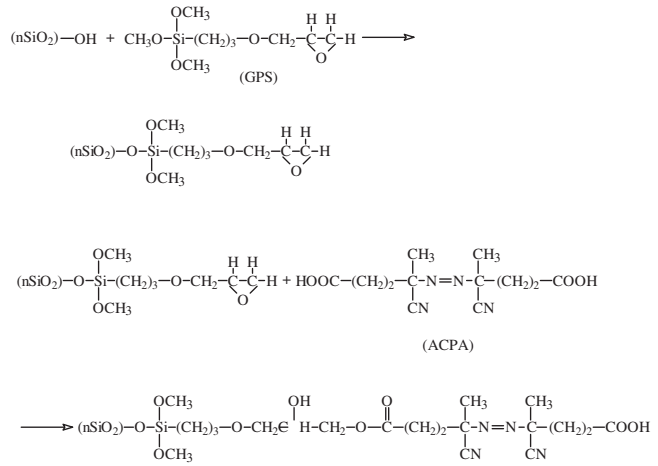
Figure 14. Introduction of peroxyester (POE) groups onto silica. Reprinted with permission from [84], N. Tsubokawa et al., *Polym. Bull.* 31, 456 (1993). © 1993, Springer-Verlag.

an amide bond established between ACPA and an amine-bearing silane coupling agent on silica, an azo initiator can also be immobilized on silica substrates [89].

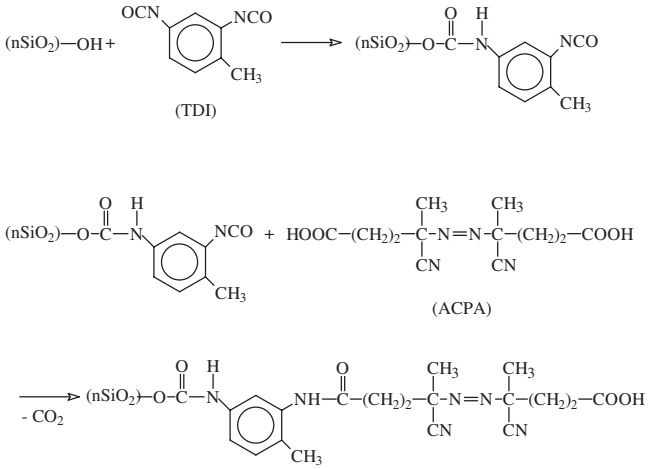
Various vinyl monomers, such as methyl methacrylate, styrene, acrylonitrile, and *N*-vinylcarbazole, were found to be effectively grafted to inorganic particles (silica (16 nm), titanium oxide (120 nm), and ferrite (15 nm)) initiated by the attached azo groups. Needless to say, these polymer-grafted nanoparticles produced stable dispersion in organic solvents. Furthermore, the dispersion of polymer-grafted ferrite in organic solvent was found to behave like a magnetic fluid.

In fact, the azo groups introduced onto inorganic particles can also initiate graft polymerization by a photoirradiation technique [88]. Furthermore, the percentage of grafting by photografting polymerization is much higher than that initiated by thermal decomposition [86], indicating that the blocking effect of surface radicals by the grafted polymer chain is considerably reduced. In the process of photopolymerization at room temperature, the propagation rate of polymer radicals is smaller than that during the thermal polymerization at 70 °C. The surface radicals and the parted free radicals have the same chance to react with the monomers. However, in the thermal polymerization system, the formation of ungrafted polymer by free radicals preferentially proceeds with the progress of polymerization.

Tsubokawa et al. revealed that a redox system consisting of ceric ion and silica particles carrying reducing groups, such as alcoholic hydroxyl, amino, and mercapto, is also capable of initiating a radical polymerization (Fig. 15) [90]. The introduction of alcoholic hydroxyl groups onto



Scheme 4



Scheme 5

the surface can be achieved by the treatment of the silica with 3-glycidoxypropyltrimethoxysilane (GPS) under acidic conditions in water. Amino and mercapto groups are introduced by the reaction of surface silanol groups with 3-aminopropyltriethoxysilane (APS) and 3-mercaptopropyltrimethoxysilane (MPS), respectively (Fig. 16). The initiating ability of the redox system changes with the reducing activity of the reducing groups on the silica surface in the following order: alcoholic hydroxyl < amino < mercapto.

Based on the above-mentioned redox systems, polyacrylamide (PAAM) was grafted to silica particles, and the grafting percentage reached about 25% [90]. It was found that both carbon and oxygen radicals can initiate the polymerization (Fig. 16). Since PAAM chains grafted to the silica surfaces interfere with the aggregation of the nanoparticles, the dispersibility of the modified silica in water was remarkably improved.

Shirai et al. proved that haloalkyl groups introduced onto ultrafine inorganic particles can serve as a coinitiator for free radical grafting polymerization based on the interaction with transition metal carbonyls [91]. In their work trichloroacetyl groups were first introduced onto the surface of TiO_2/Si-R-OH through the reaction of the surface alcoholic hydroxyl groups with trichloroacetyl isocyanate. The radical grafting polymerization of vinyl monomers was successively initiated by a system consisting of $Mo(CO)_6$ and TiO_2/Si-R-$COCCl_3$. The percentage of grafting and grafting efficiency of the graft polymerization initiated by this coinitiation system are

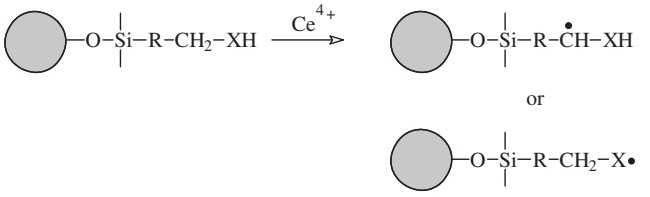


Figure 15. Generation of radicals on the silica particles carrying reducing groups. Reprinted with permission from [89], N. Tsubokawa et al., *Polym. J.* 21, 475 (1989). © 1989, Society of Polymer Science, Japan.

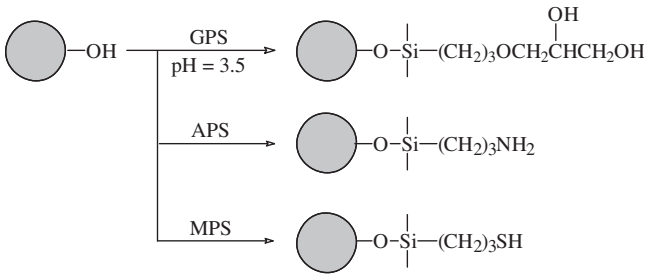


Figure 16. Introduction of reducing groups onto silica particles. Reprinted with permission from [90], N. Tsubokawa et al., *Polym. J.* 21, 475 (1989). © 1989, Society of Polymer Science, Japan.

much higher than those initiated by azo groups, because of the absence of fragment radicals.

It is worth noting that a proper selection of polymer species, along with a control of molecular weight and density of the grafted chains, is needed in all cases to obtain desired properties of polymer-grafted particles. Therefore, factors including types of initiator, monomer concentration, polymerization temperature, etc. should be considered comprehensively.

For example, the molecular weight of PS grafted to silica obtained from radical graft polymerization was found to be much higher than that from cationic graft polymerization (Table 3). The number of grafted PS in the radical polymerization, however, is much less than that in the cationic polymerization. The difference can be attributed to the initiation efficiency. That is, the initiation efficiency was lower in the former case because of the blocking effect of the previously formed polymer chains. In the cationic polymerization system, in contrast, the initiating efficiency was quite high but the chain lengths of the grafted polymer were rather short, as a result of chain transfer reactions of the growing polymer cation. On the other hand, the molecular weight of PS grafted to silica obtained from radical grafting polymerization initiated by surface azo and peroxyester groups decreased with decreasing monomer concentration and polymerization temperature. The former was assumed to be related to the reduction of the polymerization rate, whereas the latter was assumed to be related to the decrease in the number of propagating chains from the surface with a drop in polymerization temperature (Fig. 17) [92]. By means of a chain transfer agent, the molecular weight of grafted PS can also be controlled to some extent without lowering the number of grafted chains (Fig. 18).

Among various polymerization methods, a controlled/"living" polymerization technique appears to be optimal in

Table 3. Molecular weight and number of grafted PS on silica surface.

Ultrafine particles	M_n[a] ($\times 10^4$)	M_w[b] ($\times 10^4$)	M_w/M_n	Grafting number (μmol/g)
Silica-POE	55.9	120.0	2.1	0.74
Silica-APC	0.27	0.43	1.6	238.8

[a] M_n, number-averaged molecular weight.
[b] M_w, weight-averaged molecular weight.
Source: Reprinted with permission from [84], N. Tsubokawa et al., *Polym. Bull.* 31, 456 (1993). © 1993, Springer-Verlag.

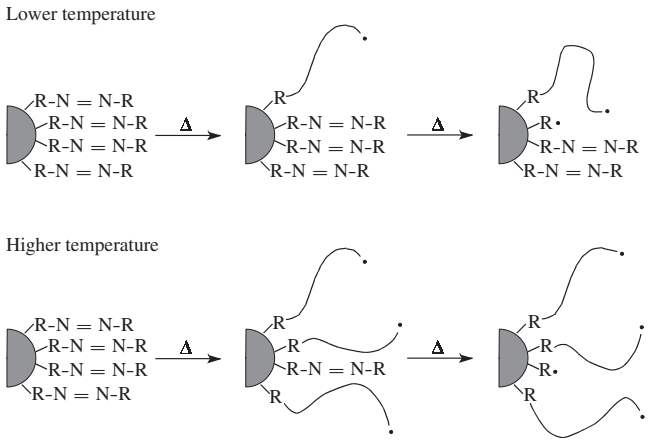


Figure 17. Effect of polymerization temperature on the number of propagating chains on particles' surfaces. Reprinted with permission from [92], N. Tsubokawa et al., *Colloid Polym. Sci.* 273, 1049 (1995). © 1995, Springer-Verlag.

controlling the molecular weight, molecular weight distribution, and structure of the grafting polymer [93]. Recently von Werne and Patten reported a technique of atom transfer radical polymerization (ATRP) for conducting controlled/"living" radical polymerizations from the surface of silica nanoparticles [94]. For this research, two initiators were prepared: (3-(2-bromopropionyl)propyl) dimethylethoxysilane (BPDS), in which the secondary α-bromoester could function as a styrene or acrylate ATRP initiator, and (3-(2-bromoisobutyryl)propyl) dimethylethoxysilane (BIDS), in which the tertiary α-bromoester could serve as a methacrylate ATRP initiator. Two sets of silica particles were used in this study, 75-nm and 300-nm diameter particles, with different surface area-to-volume ratios (0.08 vs 0.02, respectively). A monolayer of the ATRP initiator was deposited on the silica nanoparticle surface by heating the siloxane initiator and nanoparticles in tetrahydrofuran

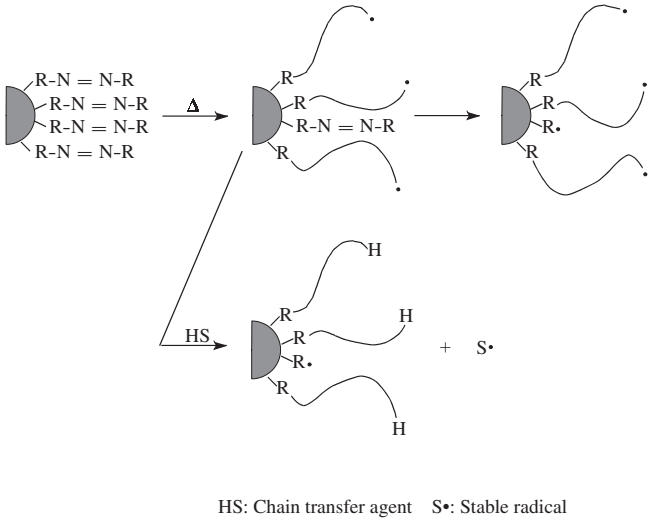


Figure 18. Effect of chain transfer agent on the grafting polymerization on particles' surfaces. Reprinted with permission from [92], N. Tsubokawa et al., *Colloid Polym. Sci.* 273, 1049 (1995). © 1995, Springer-Verlag.

(THF) at 80 °C. The standard polymerizations of styrene and methyl methacrylate (MMA) were conducted by heating the initiator-functionalized silica with monomer, copper(I) bromide, ligand, and solvent at a temperature of 90–110 °C (typical conditions: temperature 110 °C, solvent *p*-xylene, $[\text{styrene}]_o = 8.73$ M, $[\text{CuBr}]_o = 61.6$ mM, $[\text{dNbipy}]_o = 116$ mM, and 1.00 g of SiO_2/initator). For the smaller (75 nm) silica nanoparticles, the grafting molecules of PS exhibited good molecular weight control (molecular weight polydispersity index <1.25), whereas PMMA grafting chains had higher molecular weight polydispersity index of about 2.0 under the same reaction conditions. When a small amount of free initiator was added to the polymerization solution, better molecular weight control was observed. Thus, the above difference should be induced by different termination modes during polymerization processes for PS and PMMA. In the case of larger (300 nm) silica nanoparticles, neither PS nor PMMA grafting molecules exhibited molecular weight control. This lack of control was ascribed to the very high initial monomer-to-initiator ratio. It should be noted that only about 25% of the initiator present on the nanoparticle surface can initiate the growth of chains because of the sterically blocking effect of growing chains on the accessing of the catalyst to the neighboring initiation sites. As a result, the experimental molecular weights were a factor of ~4 higher than would be expected based upon the polymerization conversion factored with the ratio of monomer to nanoparticle initiator content.

Tsubokawa and Yoshikawa [95] prepared silica particles grafted to various polymers with well-defined molecular weight and structure by termination of living polymer cations with the amino groups already introduced on the particles. The introduction of amino or *N*-phenylamino groups onto the silica surface was made by the treatment of silica with γ-aminopropyltriethoxysilane or *N*-phenyl-γ-aminopropyltriethoxysilane. It was found that these amino groups on silica can readily react with living poly(isobutyl vinyl ether) and poly(2-methyl-2-oxazoline). The percentages of amino and *N*-phenylamino groups used for the grafting reaction with living poly(isobutyl vinyl ether) were only 4.4% and 4.9%, respectively. This means that the surface amino groups were blocked by the neighboring grafted polymer chains.

Liu and co-workers [96] showed the feasibility of using diisocyanate as a coupling agent to introduce poly(ethylene glycol) (PEG) (weight averaged molecular weight = 1500) to the surface of nanoapatite particles. With the use of a monoisocyanate isocyanatoethyl methacrylate, the reactivity of the surface hydroxyl groups of nanoapatite was proved. The percentage of grafted polymer was about 20% by weight. The amount of hydroxyl groups of nanoapatite was decreased by 7.7% after a surface grafting reaction. Green et al. [97] found that the grafting of PEG to alumina powder (0.5 μm) can be obtained simply by heating the dispersion of alumina powder in a pure PEG melt. The amount of the grafted polymer is strongly dependent on temperature. Grafting from a solution of the polymer, instead of the polymer melt, was also found to be effective, and a poor dispersion medium is profitable to the grafting reaction.

2.2.3. Grafting Polymerization by Irradiation

The above-described methods of chemical grafting possess advantages in controlling the development and molecular weight of the grafting polymers. However, there are also some evident disadvantages. For example, most of the grafting reactions have to be performed in solvent, which causes environmental pollution, especially in the case of large-scale production. Furthermore, the chemical techniques are quite complicated in certain cases because more than two reaction steps and a posttreatment procedure are often needed. Thus, a simpler approach to grafting by irradiation became a focus of interest of several researchers.

Grafting of polymers to inorganic particles can be realized by γ-irradiation, but only a few papers have been concerned with nanoparticles. Fukano and Kageyama [98–105] published a series of works on irradiation-induced grafting polymerization of styrene and methyl methacrylate absorbed to several kinds of nanoparticles, including silica gel (SiO_2; specific surface area 480 m^2/g), white carbon (SiO_2, 180 m^2/g), silicic acid anhydride (SiO_2, 380 m^2/g), zeolite ($Na_2Al_2Si_3O_{10} \cdot nH_2O$, 116 m^2/g), and activated alumina (Al_2O_3, 153 m^2/g). They found that the amount of grafted polymer varies with the type of inorganic substrate, with a maximum of about 30%. In the case of substrates that are chemically identical, the amount of grafted polymer is proportional to the specific surface area of the particles. In the case of substrates with the same specific surface area, inorganic substances that contain aluminum as a component element are grafted more easily than those consisting of SiO_2 alone. Gel permeation chromatography (GPC) measurements indicated that both the grafted polymer and the homopolymer possess low- and high-molecular-weight chains. By studying the effect of inhibitors (*p*-benzoquinone and ammonia) on the ratio between the low- and the high-molecular-weight portions of the grafted polymer and the homopolymer, mechanisms involved in irradiation-induced polymerization were concluded. It was believed that the high-molecular-weight materials appear to have been generated by radical polymerization, whereas the low-molecular-weight polymers are formed as a result of cationic (for styrene-silica) and anionic (for methyl methacrylate-silica) polymerizations. Further studies revealed that the relative yields of the low- and high-molecular-weight parts of the grafted polymer and the homopolymer are a function of dose rate and irradiation temperature.

It should be noted that a postirradiation effect exists in certain cases when methyl methacrylate is grafted to MgO by simultaneous irradiation methods [106]. Such an effect is characterized by the appearance of long-lived radicals or ions, which might lead to a continuous polymerization, even when the system is away from the irradiation field. Certain measures should thus be taken to eliminate the remaining radicals and ions to stabilize the microstructure of the grafted particles.

Fukano and Kageyama have also tried the preirradiation grafting polymerization of styrene to silica gel surfaces [105]. It was found that both radical and cationic polymerizations took place to form grafting polymer and homopolymer in a way similar to that of the case of simultaneous irradiation.

To obtain a higher degree of grafting, Fukano and Kageyama conducted an irradiation grafting process under

high vacuum conditions (lower than 10^{-4} Hg) [98–105]. Actually, this is less convenient for industrial application. Therefore, Rong et al. developed another method in which strict irradiation grafting conditions, such as high vacuum and purification of monomers, usually required in the case of conventional particles, are no longer necessary [107]. Various polymers were grafted to the surface of silica nanoparticles through simultaneous irradiation (Table 4). The molecular weights (M_n and M_w) of both the grafted and the homopolymerized polystyrene obtained by irradiation with or without solvent are shown in Table 5. The higher M_n of the homopolymer, relative to the grafting polymer, could be ascribed to the stronger motional ability of nanoparticles compared with conventional particles, which made the chain termination between the two radicals easier. The results also showed a relative wide molecular weight distribution of PS in the system with the solvent; this is due to the fact that chain transfer from the radical to the solvent can easily take place. FTIR and ESCA were adopted to characterize the interaction between the nanoparticles and the grafted PS, demonstrating that PS is chemically connected with SiO_2 through Si-O-C and Si-C bonds.

2.3. Additional Remarks

Based on the above introduction about surface treatments for nanosized inorganic particles, it can be concluded that there is no evident difference between the treatments of microparticles and nanoparticles. However, some issues relating to the practical application of the modification approaches with respect to the specific structure and properties of nanoparticles should be emphasized as follows:

- A nanoparticle assembly usually possesses a hierarchy structure, starting from primary particles, via aggregates to agglomerates (in which aggregates are bound through hydrogen and electrostatic bonding forces). A surface modification may cause a splitting of agglomerates or aggregates to a certain extent according to the treatment conditions. However, the decomposed aggregates can get together once again when they are incorporated into a polymer matrix, because of the uneven coverage of the modifier molecules. Therefore, the eventual filler size and its distribution in the composites depend not only on the surface treatment but also on the compounding technique. Meanwhile, it should be kept in mind that primary particles in a premade nanoparticle assembly are hard to attain no matter what treatment is used.
- When compared with microparticles, nanoparticles, with their tremendous specific surface area, have a reasonably high surface activity and would lead to substantial chemical interactions, even in the case of a physical treatment. On the other hand, there are much greater difficulties encountered by surface modifiers in penetrating the tightly bonded nanoparticle agglomerates, especially for physical treatment methods such as surfactant treatment and encapsulation of polymer.
- Chemical treatment is more effective in promoting the miscibility between nanoparticles and polymers, but the penetration of nanoparticle agglomerates by reactive agents is still a problem. By comparison, infiltration of monomers followed by polymerization benefits the disintegration of the agglomerates.
- Morphologies of grafting or encapsulating polymer chains have scarcely been studied because of the limitation of characterization facilities. Research in this direction is of great importance for controlling the particle dispersion in the final matrix, because the morphology of the molecules surrounding the nanoparticles is closely related to the entanglement with the polymer chains of the matrix, especially for those nanocomposites produced by solution mixing. Therefore, the length of grafting or encapsulating chains, as well as the number and position of reactive groups, should be considered in detail to optimize the surface treatment effect.
- Most of the published works that deal with surface treatment of nanoparticles neglect the correlation between treating methods and structural details of treated nanoparticle agglomerates with their interfacial interactions, which occur when the particles are added to a polymer matrix. In other words, various surface treatments developed up to now only focus on treatment methods themselves. Their application in preparing nanoparticle-filled polymer composites has not been studied deeply. Therefore, in the following sections, microstructural effects of treated nanoparticles on the properties of the final nanocomposites

Table 4. Grafting details of nano-SiO_2 grafted with different polymers.

Grafting polymers	PS	PBA[a]	PVA[b]	PEA[c]	PMMA	PMA[d]
Percentage grafting[e] (%)	3.64	3.32	2.82	1.73	1.85	2.16
Homopolymer fraction[f] (%)	16.3	15.4	13.7	12.3	15.5	15.4

[a] Polybutyl acrylate.
[b] Polyvinyl acetate.
[c] Polyethyl acrylate.
[d] Polymethyl acrylate.
[e] Weight of grafting polymer/weight of SiO_2.
[f] Weight of homopolymer/weight of SiO_2.

Irradiation dose = 10 Mrad; weight ratio of monomer/SiO_2= 20/100. All of the systems used acetone as a solvent when they were irradiated, except for the methyl acrylic acid/SiO_2 system, in which ethanol was used as a solvent.

Source: Reprinted with permission from [107], M. Z. Rong et al., *Polymer* 42, 167 (2001). © 2001, Elsevier Science.

Table 5. Molecular weights and their distributions of grafting PS and homopolymerized PS in the system of nano-SiO_2 grafted with PS.

Irradiation conditions	Materials	M_w ($\times 10^4$)	M_n ($\times 10^4$)	M_w/M_n
Without solvent	Grafting PS	3.5	1.7	2.1
	Homopolymerized PS	7.3	4.3	1.7
With solvent	Grafting PS	5.3	2.2	2.4
	Homopolymerized PS	2.1	6.3	3.3

Note: Irradiation dose = 10 Mrad.

Source: Reprinted with permission from [107], M. Z. Rong et al., *Polymer* 42, 167 (2001). © 2001, Elsevier Science.

will be discussed mainly from the processing point of view.

3. COMPOUNDING AND MECHANICAL PERFORMANCE OF POLYMER NANOCOMPOSITES

3.1. Fundamentals of Agglomerate Dispersion in Polymers

Dispersive mixing in polymer-based particulate composites has an important technical meaning. The state of dispersion usually refers to the amount of unagglomerated particulates that are present in a compound. The number and nature of the dispersed solids have been shown to greatly influence the mechanical behavior of the composites. The appearance of nondispersed or reagglomerated clusters is detrimental to the uniformity of the ultimate properties of the products [108]. As this chapter deals with ready-made nonlayered nanoparticles that are used in the form of agglomerates, problems related to agglomeration and deagglomeration should be considered throughout the discussion. In particular, understanding the dispersion process and the relationships between processing conditions and quality and extent of dispersion are of practical interest.

3.1.1. Variation of Agglomerates during Mixing

During the compounding operation, the dispersion is generally achieved in four steps: incorporation, wetting, agglomerate breakup, and aggregate spatial distribution in a polymer matrix [109]. It must be ensured that the particulate fillers can be wetted and distributed before they can agglomerate. In the field of solid-liquid dispersion, a series of works have been devoted to the breakup of agglomerates. Agglomerate breakup occurs when hydrodynamic forces induced by the flow overcome the cohesion forces mentioned in the previous chapter. Consequently, the mixing intensity needed is strongly dependent on the affinity between the polymer matrix and the additives.

The mechanism of agglomerate dispersion in various media has been studied by several researchers. Two major types of cluster breakup have been identified; in particular, Parker et al. proposed the occurrence of (1) large-scale splitting, also called the rupture process, and (2) fine particle erosion [110]. Rwei et al. suggested the existence of critical conditions for the different dispersion mechanisms [111, 112]. Erosion, which is initiated at lower shear stresses, is characterized by a continuous detachment of small fragments from the outer cluster surface. The rupture mechanism is characterized by an abrupt splitting of the agglomerate into a small number of large aggregates; it tends to take place at relatively high shear stresses. Analysis of the evolution of carbon black agglomerates in the melts of PS and high-density polyethylene (HDPE) indicated that a coarse rupturing of the agglomerates occurred during the early stages of the dispersion. This was followed by a more gradual erosion of small aggregates from large fragments [113]. Pandya and Spielman investigated the rupture process of cohesive agglomerates and found that fragment size and configuration were affected by the shape and internal structure of the parent agglomerate [114]. Erosion of cohesionless clusters of spherical particles was studied by Mason et al.; they also modeled the effect of shear stress and particle size on the erosion rate [115, 116]. Shiga and Furuta suggested that the dominant mechanism of carbon black agglomerate dispersion in elastomers is the scraping of individual constituent particles from the surface of the agglomerates [117]. Most of the models based on experimental investigations show that the variation of the agglomerate size as a function of time follows a first-order kinetic law regardless of the matrix viscosities [112]:

$$\ln[R(t)/R_o] = Kt\dot{\gamma} \tag{3}$$

where $R(t)$ is the radius of the cluster at time t, R_o is the initial cluster radius, K is the erosion rate constant, $\dot{\gamma}$ is the applied shear rate, and $t\dot{\gamma}$ is a dimensionless erosion time (Fig. 19). It is believed that K is dependent on the flow geometry, the applied shear stress, and the cohesive strength of the agglomerate.

For a short erosion time, Eq. (3) leads to

$$[R_o - R(t)]/R_o = Kt\dot{\gamma} \tag{4}$$

In this case, the erosion kinetics was found to be independent of both the agglomerate size and the applied shear rate. When interfacial interactions between particulate agglomerates and the polymer matrix are taken into consideration, variation of the agglomerate size can be described by

$$[R_o - R(t)]/R_o = K'W_{ab}\mu\dot{\gamma}t/\tau_c \tag{5}$$

where K' is the erosion rate constant corrected for the work of adhesion, τ_c is the cohesive strength of the agglomerate, and μ is the fluid viscosity [118]. This model has been confirmed by the radius reduction of ultrafine TiO_2 agglomerates in linear low-density polyethylene (LLDPE) melts with shearing time.

Since particulate agglomerates have a permeable structure, they can be subjected to a fluid infiltration. The infiltration of a liquid polymer to a porous agglomerate

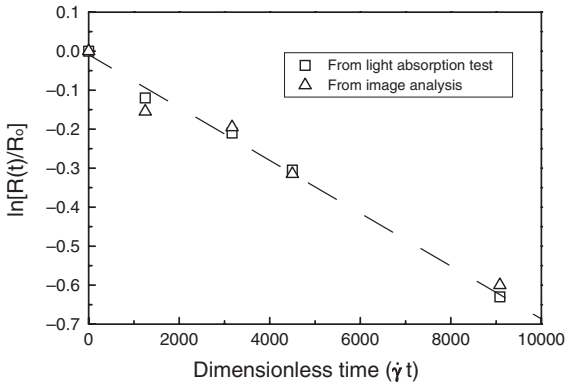


Figure 19. Carbon black agglomerate size reduction in PS melt as a function of dimensionless shearing time under a shear stress of 19,500 Pa. Reprinted with permission from [113], S. P. Rwei et al., *Polym. Eng. Sci.* 32, 130 (1992). © 1992, Society of Plastics Engineering.

corresponds to a transition from a dry agglomerate (pores filled with air) to a wetted one (pores filled with polymer melt). This process is driven by the capillary pressure (pressure difference across the interface between the two filling media) and the hydrostatic pressure, although the latter can be shown to be negligible compared with the former. Liquids characterized by a contact angle on particulate fillers of less than a critical value can infiltrate the agglomerates, provided the air displaced from the pores is allowed to escape. Otherwise, the air pressure inside the agglomerate rises to a point where it counterbalances the capillary pressure, and consequently the infiltration stops.

The presence of a liquid inside an agglomerate affects its cohesiveness. The creation of liquid bridges leads to a substantial increase in the tensile strength of moist agglomerates [119]. Calcium carbonate agglomerates, for example, were found to be strengthened by matrix infiltration during dispersion in polymer melts under simple shear flow conditions. Hence assessing the kinetics of matrix infiltration constitutes a valuable tool in the analysis of the dispersion behavior of fine particle agglomerates in molten polymers.

The role of matrix infiltration in the hydrodynamic dispersion of particulate clusters was highlighted during the study of the dispersion of carbon black agglomerates in liquid polymers. Yamada and co-workers investigated the dispersion mechanisms of carbon black agglomerates in simple shear flows. They found that the agglomerate erosion kinetics could be interpreted in terms of the degree of matrix infiltration and agglomerate permeability [120, 121]. The agglomerate dispersion behavior could be correlated with the ratio of the depth of the infiltrated layer to the square root of the agglomerate permeability, a characteristic length related to the thickness of the agglomerate region that can be drained by viscous flow.

According to the model proposed by Yamada et al. for the kinetics of agglomerate infiltration [122], it is assumed that matrix infiltration in spherical agglomerates occurs symmetrically, leading to an infiltrated outer shell of constant thickness and a dry core. The principle of a spherically symmetric infiltration has been confirmed experimentally by Bohin et al. by phase-contrast microscopy [123]. The kinetics of infiltration of $CaCO_3$ agglomerates by liquid polymers can be well described by the theoretical relationship based on Darcy's law [124]. Reducing the agglomerate radius, polymer viscosity, and/or agglomerate density leads to higher infiltration kinetics. A stearic acid surface treatment, for example, was observed to have a dramatic influence on the latter. Infiltration was much slower for the surface-coated $CaCO_3$ (80 nm) than for the untreated powder when other parameters were kept identical. It was suspected that the surface treatment could have affected the morphology and the packing characteristics of the nanoparticles.

3.1.2. Mechanical Deagglomeration

Intimate contacts (up to almost the molecular level) of the nanoparticles with the polymeric matrix are a prerequisite for the design of well-functioning nanocomposites. To satisfy this condition, a thorough homogenization of the components must be obtained. Winkler et al. [125] developed a theory for the deagglomeration of ultrafine particle clusters in fluid systems, using the mechanical forces in a dispersion machinery. They showed that the probability that filler particles are dispersed in a dispersion medium, ρ_T, is the product of two separate, partial probabilities: ρ_t, the probability with which the agglomerates encounter a potentially active dispersion site, and ρ_ε, the probability that enough energy per unit volume (energy density) is available to overcome the forces holding the agglomerates together. By applying their theory to dispersion experiments carried out in a bead mill and a high-speed impeller [126], the effective volumes per second (in which a dispersion of the agglomerates could take place) were found to be on the order of 1% of the total volume, and the transfer efficiency of the energy into the agglomerates was approximately 1% as well. According to these findings, it was suggested that the key to success for dispersing particles is to guide their agglomerates in a constrained path through well-defined zones of suitable energy densities. In addition, Winkler and Dulog found that the deagglomeration in most dispersion machinery is governed by the shear stress within the mill base [127]. This means that smearing-type mechanisms are far more important than smashing-type stresses for dispersions.

From a practical point of view, compounding polymers in their molten state with additives is generally achieved by various devices, like two-roll mills, internal mixers, single-screw extruders, and twin-screw extruders [128]. Two-roll mills and internal mixers are not suitable for production on an industrial scale. Single-screw extruders cannot provide sufficient homogenization. Corotating twin-screw extruders are therefore the most popular machines because of their continuous operation and high productivity. Both the screw and the barrel of the extruders are constructed from segments and can be assembled to produce the necessary conditions for efficient homogenization.

Fundamental dispersion studies have identified some useful rules for optimum compounder performance, for example [129–131]:

- Since dispersion is a phase reduction process, it pays to generate a well-mixed feed in the feeder train or feed section of the compounder.
- High stresses applied during melting are important and friction dependent, and care should be taken to reserve lubricating fluids, powders, and low melting solids for post-addition.
- Compatibilizers reduce the interfacial tension and result in dramatically improved dispersions.
- Solid particles in polymer melts break up like a fatigue failure, which accounts for the fact that there must be a minimum residence time established to achieve dispersion in addition to a minimum shear stress. Above the minimum residence time and minimum shear stress, the shear and time values are interchangeable in achieving the same dispersion effect.
- Shear yield stress of particle filled polymer melts increases with decreasing particle size.

When a twin-screw extruder is used, the effects of extrusion processing conditions on the final state of dispersion should be known in advance, so as to control the formation of agglomerates. Although most of the available literature is

actually based on the approaches developed for microcomposites, the original principles stated above might also be applicable to nanocomposites. This is due to the facts that (a) the size of the nanoparticle agglomerates also falls in to the micrometer regime, and (b) the initial particle size gives only slight differences in dispersion, as demonstrated by Ess and Hornsby in their paper dealing with $CaCO_3$ particles varying in diameter from 80 nm to 3 μm [132].

Ess and Hornsby discussed the dispersive mixture quality of $CaCO_3$-filled PP and PA 66 composites compounded by an intermeshing corotating twin-screw extruder. For the system under investigation, the zone of polymer fusion at the end of the first stage of the screw profile was identified as having a crucial influence on enhancing the filler dispersion because of the high levels of shear developed in this region. Furthermore, there was some evidence of filler agglomeration due to compactive forces in the solid bed formed before polymer melting. Compounding experiments, using pellets of $CaCO_3$ preconditioned under various pressure, temperature, and moisture conditions, indicated that too high values of these quantities gave a marked reduction in the degree of filler dispersion. These results therefore highlight the importance of controlling these variables (for example, in premixing operations) for an optimized filler dispersion.

Recently, Bories et al. made a study of this aspect by using polypropylene (PP) and $CaCO_3$ (0.7 μm) as the experimental materials [133]. They found that reducing the extrusion flow rate or increasing the rotation speed and barrel temperature can yield improved dispersion. Furthermore, the main concerns of their work were focused on twin-screw extruder design and its relation to particle agglomerate dispersion. Four different screw geometries were employed; each one had two separate kneading sections, referred to as the melting and mixing sections (Table 6). To evaluate the conditions that minimize the relative agglomerate fraction, the dispersion efficiency of the four different screw configurations was compared with the use of either hopper feeding or mid-extruder feeding of $CaCO_3$. As shown in Figure 20a, where the screw configurations have a common severe melting section and differ in the mixing section, better dispersions are achieved when PP and $CaCO_3$ are fed together at the hopper. It is possible that a distribution of the mineral can be achieved even during solid conveying before PP melting, leading to low levels of agglomeration. When feeding takes place at the primary hopper, the state of dispersion does not change significantly whether severe or gentle mixing sections are used downstream, indicating that the agglomeration/dispersion process is completed after the melting zone. The conclusions are quite different when feeding of $CaCO_3$ takes place at the mid-extruder. In this case, significantly higher agglomerate volume fractions are observed for the severe mixing zone, which includes a reverse element that is used to fill the mixing zone. This may cause additional agglomeration, since large undistributed mineral pockets can be rapidly compressed into agglomerates in this section.

Table 6. Summary of screw configuration severity.

Screw configuration	Melting section (first mixing section)	Mixing section (second mixing section)
1	+	−
2	+	+
3	−	+
4	−	−

Note: + and − represent severe and gentle, respectively.
Source: Reprinted with permission from [133], M. Bories et al., *Int. Polym. Process*. 14, 234 (1999). © 1999 Carl Hanser Verlag.

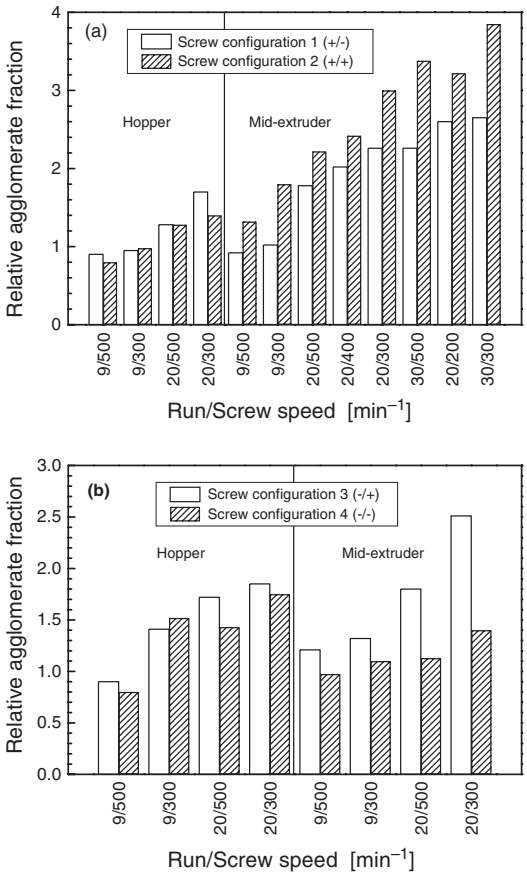


Figure 20. Relative agglomerate volume fraction obtained with different screw configurations for a 60 wt.% $CaCO_3$/PP composite (barrel temperature 200 °C). Reprinted with permission from [133], M. Bories et al., *Int. Polym. Process.* 14, 234 (1999). © 1999, Carl Hanser Verlag.

When the screws have a common gentle melting section and differ in the second mixing zone severity (Fig. 20b), the best results are obtained by mid-extruder feeding. For both hopper and mid-extruder feeding, the best improvements are achieved with configuration 4, which has gentle melting and mixing zones. The use of a severe mixing section is detrimental for the dispersion in both configurations, leading to higher agglomeration levels. The effect of the second mixing zone is also observed, even when feeding of the $CaCO_3$ takes place in the primary hopper. This indicates that the dispersion process is not completed at the end of the melting zone, contrary to the observations with the severe melting section. The PP, being only partially molten in the melting zone, is still highly viscous at the mid-extruder position. This seems to be a good compromise. The polymer has sufficient fluidity to help in the wetting of the mineral while it is still viscous enough to generate high hydrodynamic stresses. Best overall results are thus obtained when the mineral is fed at mid-extruder with the gentlest configuration.

3.2. Polymer Nanocomposites Produced Through Different Mixing Techniques

3.2.1. Melt Mixing

Melt mixing is most commonly used because of its energy efficiency, operability, and environmental containment. The proven mechanisms for transporting viscous melts include drag and ram flow. Phase homogenization can be significantly improved through reorientation and gap flow in the mixer. The melt mixing dispersion process is controlled by the shear viscosity and elasticity ratios of the components and the deformation field generated by the operating condition and the mixer geometry. Although the publications dealing with direct incorporation of nonlayered nanoparticles into polymers are relatively few in comparison with those using intercalation of layered solids, most of the available papers in this field are based on the melt mixing technique.

Huang et al. mixed SiC/Si_3N_4 particles (20 nm) with low-density polyethylene (LDPE) with the use of a two-roll mill at 120 °C for 6–10 min, followed by compression molding at 170 °C under 70 MPa [134]. The nanoparticles were pretreated with a titanate coupling agent. For comparison, microscale $CaCO_3$ (1–3 μm)- and TiO_2 (0.3–0.8 μm)-filled LDPE composites were made in a similar way. Mechanical testing of the materials indicated that both impact and tensile strengths of LDPE were doubled when 5 wt.% nano-SiC/Si_3N_4 was incorporated, whereas neither $CaCO_3$ nor TiO_2 could provide any reinforcing effect. According to the melt flow rate measurements, a drastic drop was perceivable at a nano-SiC/Si_3N_4 loading of 5 wt.% (Fig. 21). It was believed that a strong interaction between the nanoparticles and the matrix was established at the very low filler content. In a latter report of Huang's group [135], nano-SiC/Si_3N_4 was employed to improve the mechanical performance of conductive polymer composites made of carbon black (CB)-filled chlorinated polyethylene (CPE)/polyvinylchloride (PVC). The relationships between mixing time for roll milling at 172 °C and strength properties were studied carefully (Fig. 22). It was found that a mixing over 9 min gave optimum results. Furthermore,

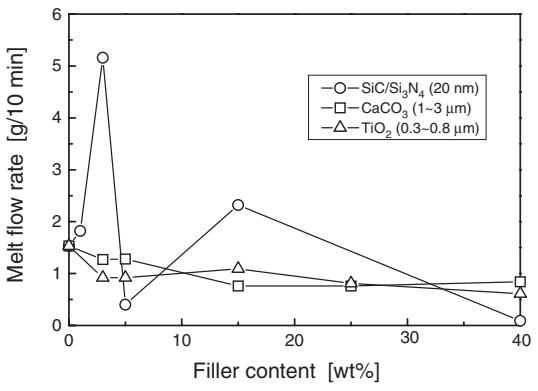


Figure 21. Melt flow rates of LDPE-based composites. Reprinted with permission from [134], R. Huang et al., *Chin. Plast. Ind.* 25, 106 (1997) (in Chinese). © 1997, Chenguang Research Institute of Chemical Industry.

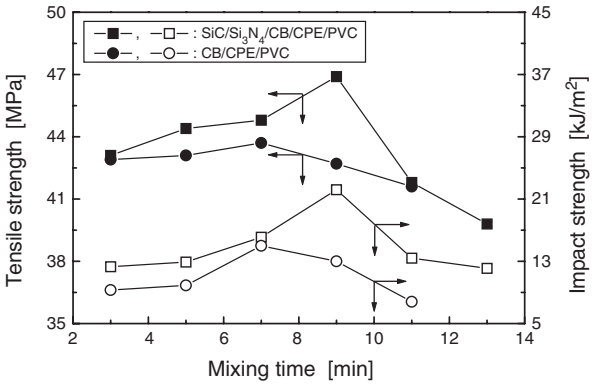


Figure 22. Tensile and impact strengths of CPE/PVC (blending ratio = 12.5/100 by weight)-based nanocomposites as a function of mixing time (SiC/Si_3N_4 content = 0.5 wt.%, CB content = 15 wt.%). Reprinted with permission from [135], Y. Su et al., *Chin. Plast.* 12, 22 (1998) (in Chinese). © 1998, Chinese Plastics Processing Industry Association.

the addition of nano-SiC/Si_3N_4 reduced the melt viscosity of CB/CPE/PVC systems. Their data are an indicative of positive hybrid effects of nanoparticles. That is, a combination of different nanoparticles (like SiC/Si_3N_4 and CB in the present system) can yield a composite with performance and processability better than those of one containing only one type of particle. It manifests the feasibility of preparing polymer nanocomposites coupled with multiple functions (e.g., flame retardant, electroconductive, etc.).

To further improve the compatibility between nanoparticles and polyolefines, Xu et al. introduced into the treatment of nanoparticles a macromolecular coupling agent that contains hydrophobic carbon chains much longer than those found in conventional coupling agents [136]. In their work, $CaCO_3$ (50 nm) and HDPE were employed as the basic materials, and P403 macromolecular coupling agent and NDZ-101 titanate coupling agent were added to the nanoparticles for surface pretreatment, respectively. Rheological investigations of the composites at 200 °C indicated that the two types of coupling agents helped to reduce the apparent viscosity, especially under low shear rates. When the shear rate increased to 500 s^{-1} and above, the rheological behavior of the filled systems approached that of the neat HDPE. Considering that plastics are usually extruded under a shear rate of 10^2 to 10^3 s^{-1} and injection is molded under 10^3 to 10^5 s^{-1}, it was expected that the incorporation of the treated nano-$CaCO_3$ would not change the original processability of HDPE. With respect to the tensile properties [137], the addition of P403-treated $CaCO_3$ prominently improved the elongation to break of the composites, in comparison with NDZ-101-treated $CaCO_3$/HDPE (Table 7). However, the nanoparticles treated with both coupling agents failed to enhance the strength of the matrix polymer.

Guo et al. compared the effect of various silane coupling agents on the melt viscosity of SiO_2 (20 nm)-filled PMMA composites with the use of a Haake rheometer (temperature 210 °C, rotor speed 60 rpm) [138]. As can be evaluated by the maximum torques and equilibrium torques shown in Table 8, the composites with the addition of A-174-treated SiO_2 had the highest flowability. Evidently, the compatibility between methacryloxypropyl segments of

Table 7. Percentage elongation to break of nano-$CaCO_3$/HDPE composites.

Filler content (wt.%)	0	4	7	10	13	16	22	26	30
Untreated $CaCO_3$/HDPE	600	500	460	590	100	200	300	370	—
NDZ-101[a]-treated $CaCO_3$/HDPE	600	>600	>600	>600	>670	>660	>550	>420	>400
P403[b]-treated $CaCO_3$/HDPE	600	Unbroken	Unbroken	Unbroken	>670	>660	>600	630	470

[a] Titanate coupling agent.
[b] Macromolecular coupling agent (molecular weight ≈ 10,000) with the copolymer of maleic anhydride and vinyl monomer as the backbone chains and acrylic polymer as the side chains.
Source: Reprinted with permission from [136], W. Xu et al., *Chin. Plast.* 13, 25 (1999) (in Chinese). © 1999, Chinese Plastics Processing Industry Association.

A-174 and PMMA should be responsible for this. It is worth noting that the dosage of the coupling agent should not exceed a certain limit, otherwise the excess silane would act as an internal plasticizer, leading to reduced mechanical performance of the composites (Fig. 23). In the case of micrometer particulate-filled polymer composites, the optimum coupling agent/filler ratio can be estimated in terms of the so-called monomolecular layer approach [139]. It assumes an even coating of a monomolecular layer of a coupling agent on each particle. However, the experimental data of polymer nanocomposites indicated that the calculated optimum dosages are always noticeably higher than the actual ones [140, 141], manifesting the invalidity of the above theory on a nanoscale. Since the content of nanoparticles necessary to cause a consistent drop of viscosity of the polymer has to be higher than that usually adopted for coarse particulates, because of the high surface area of the nanoparticles, new rules should be determined. Levita et al. suggested a simple way to determine the appropriate amount of coupling agent by measuring the viscosity of a dispersion of nanoparticles in mineral oil (Fig. 24) [142].

In view of the fact that the polar groups in certain thermosetting resins strongly interact with active functional groups of fillers [143, 144], such as oxidized carbon fibers, and form coating layers around the filler particles [145, 146, 147], Jana and Jain expected that the polar reactive solvents could help dispersion and prevent agglomerate formation by nanoparticles [148]. They used epoxy as a processing aid and as a dispersing agent of fumed nanosilica in producing melt-mixed blends of polyethersulfone (PES). Epoxy component was polymerized after dispersion of the particles to recover the mechanical properties. It was found that strong interaction between the polar groups of epoxy and silanol groups on the surface of fumed silica helped dispersion of the particles into PES. The viscosity and processing temperature of the PES/epoxy/silica system were significantly reduced as compared with PES and PES/silica blends. Furthermore, improvement in barrier resistance to diffusion of solvents like methylene chloride, heat deflection temperature, and impact strength over neat PES was observed.

During preparation of micrometer particle-filled polymers, mixing dispersion of the fillers is usually completed by simply adding all of the ingredients of a certain proportion into the extruder. In the case of polymer nanocomposites, however, homogeneous dispersion of the nanofillers would be hard to achieve in this way because of their high surface energy and low volume fraction in the composites. Using a twin-screw extruder, Ren et al. produced a

Table 8. Torque–time relation of nano-SiO_2/PMMA composites.[a]

	Torque (Nm)			
Time (min)	A-174[b]-treated SiO_2/PMMA	A-1100[c]-treated SiO_2/PMMA	A-151[d]-treated SiO_2/PMMA	Untreated SiO_2/PMMA
0.5	2.0	2.3	2.3	2.4
1	46.1	47.4	47.9	48.4
1.5	22.0	22.0	22.8	23.2
2	14.2	15	15.5	16.3
3	9.9	10.9	10.8	11.6
4	7.6	8.2	8.4	8.5
5	6.6	7.0	7.0	7.1
6	6.0	6.2	6.2	6.3
8	5.0	5.2	5.3	5.4
10	4.5	4.7	5.2	5.4

[a] SiO_2 content = 5wt.%.
[b] γ-Methacryloxypropyl trimethoxy silane.
[c] γ-Aminopropyltriethoxy silane.
[d] Vinyltriethoxy silane.
Source: Reprinted with permission from [138], W. Guo et al., *Chin. Plast. Ind.* 26, 10 (1998) (in Chinese). © 1998, Chenguang Research Institute of Chemical Industry.

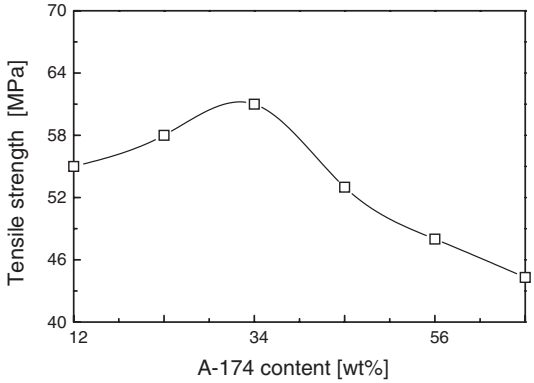


Figure 23. Tensile strength of nano-SiO_2/PMMA composites (5 wt.%) with different contents of A-174 (A-174 content is relative the amount of SiO_2). Reprinted with permission from [138], W. Guo et al., *Chin. Plast. Ind.* 26, 10 (1998) (in Chinese). © 1998, Chenguang Research Institute of Chemical Industry.

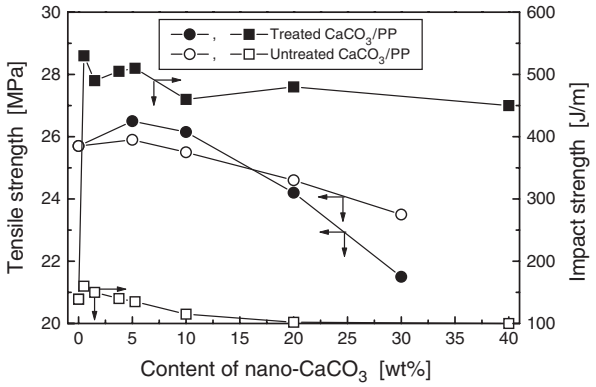


Figure 25. Tensile and notched Izod impact strengths of nano-$CaCO_3$/PP composites. Reprinted with permission from [140], X. Ren et al., *Chem. World* 41, 83 (2000) (in Chinese). © 2000, Society of Chemistry and Chemical Engineering of Shanghai.

$CaCO_3$ (80 nm)/PP nanocomposite master batch first and then compounded the master batch with the same PP to dilute the filler loading to a desired value [149]. They found that the homogeneity and the mechanical performance of the composites were much better than those of the materials prepared in one step. Recently, Zhang et al. reported similar results when studying the effect of macromolecular dispersant pretreatment of TiO_2 (20 nm), in combination with the master-batch approach, on the mechanical performance of high-impact polystyrene (HIPS) [150]. With a nano-TiO_2 content of less than 5 wt.%, not only the Young's modulus, the tensile strength, and the impact strength, but also the hardness of the composites could be enhanced. This might also act as a new solution for the improvement of the inherently poor resistance of HIPS to surface scratching. In another paper by Ren and co-workers [140], the authors evaluated the effect of a homemade phosphonate coupling agent. The treated nano-$CaCO_3$ can result in an increase in tensile and impact strength, whereas the untreated agents had nearly no reinforcing effect (Fig. 25). In addition to the effects of the surface modification of the nanoparticles, their results also showed the importance of the matrix toughness. As illustrated in Figure 26, a much more remarkable increase in impact resistance was observed in the J340-PP-based composites (a PP copolymer with higher ductility), but the same nanoparticles did not result in a similar improvement in the F401-PP homopolymer. It was suggested that the polymer to be toughened by nanoparticles should possess at least a certain toughness. This finding received experimental support from Wang et al., who mixed an elastomer with $CaCO_3$ (70 nm) into a PP matrix [151]. Both notched and unnotched impact strength values of the trinary blends were found to be much higher than those of the binary ones without the elastomer, at the same content of nanoparticles.

In most of the cases, as introduced above, people treated nanoparticles by using coupling agents or dispersants before the melt mixing procedure. However, since only the exterior nanoparticles of the agglomerates can take part in the reaction, split of the treated nanoparticle agglomerates during the subsequent compounding with polymers at high

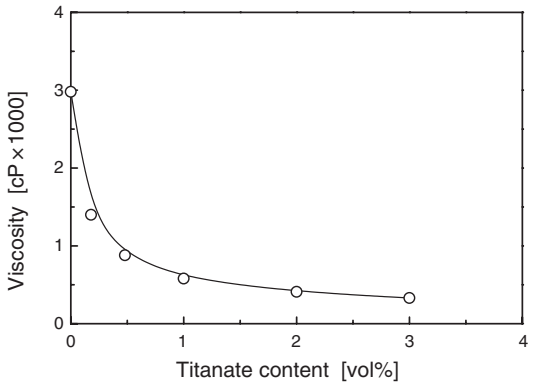


Figure 24. Effect of KR-TTS titanate on the viscosity of a dispersion of $CaCO_3$ (70 nm) in mineral oil (filler content = 24 vol.%). Reprinted with permission from [142], G. Levita et al., *Polym. Compos.* 10, 39 (1989). © 1989, Society of Plastics Engineering.

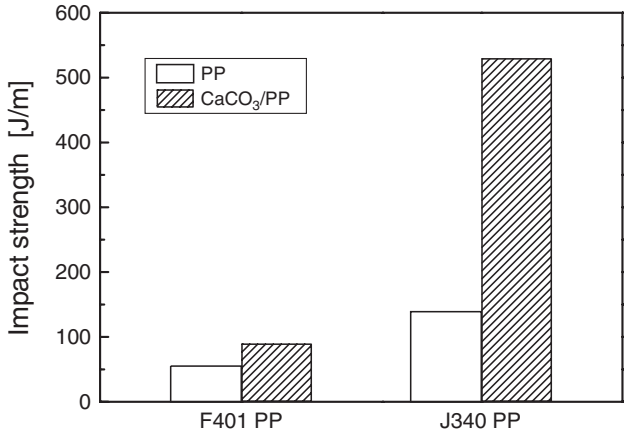


Figure 26. Comparison of notched Izod impact strengths of phosphonate coupling agent treated nano-$CaCO_3$/PP composites with different PP matrices (filler content = 0.5 wt.%). Reprinted with permission from [140], X. Ren et al., *Chem. World* 41, 83 (2000) (in Chinese). © 2000, Society of Chemistry and Chemical Engineering of Shanghai.

temperature leads to an exposure and reagglomeration of the unreacted interior particles. As a result, these newly flocculated and loosely bound nanoparticle agglomerates can act as sites of stress concentration in the composites. To solve this problem, Wu and Ye followed a different route in hopes of stabilizing the split particles [152]. They first put HDPE in a two-roll mill, preheated to 150 °C. When the polymer was melted thoroughly, both $CaCO_3$ (40 nm) and stearic acid at a desired composition were incorporated. It was expected that (a) the dispersed nanoparticles in the polymer melt can be activated *in-situ* by stearic acid, and (b) the formation of larger nanoparticle agglomerates can be prevented accordingly. Microstructural observation of the fractured surfaces of the composites proved the feasibility of their idea. That is, nearly no nanoparticle agglomerates appeared when a proper dosage of the dispersing agent was applied. For the composites containing 20 wt.% nano-$CaCO_3$ prepared in this way, both impact strength and elongation to break of HDPE were increased by more than 100%.

By inferring that it might be impossible to pursue a nanoscale dispersion of nanoparticles in polymer melts characterized by a high viscosity, Rong et al. emphasized the enhancement of the inherently friable structure of nanoparticle agglomerates besides the formation of sufficient interaction between fillers and matrix by means of an irradiation grafting technique [107, 153]. The typical pregrafting of the nanoparticles proceeded by mixing the monomers with the particles and then irradiating the mixture under ^{60}Co γ-rays at room temperature. Although the processing viscosity of the grafted nanoparticles/PP was slightly higher than that of the untreated system (Fig. 27), it was still close to that of PP, implying that the processing variables can be the same as those set for manufacture of the neat polymer. To prepare polymer composites, the treated nanoparticles were compounded with PP on a lab-scale single-screw extruder at 200 °C. The experimental results exhibited that a simultaneous enhancement in stiffness, strength, and toughness of PP can be acquired by the addition of only a small amount of modified SiO_2 (7 nm) or $CaCO_3$ (60 nm) (Figs. 28 and 29).

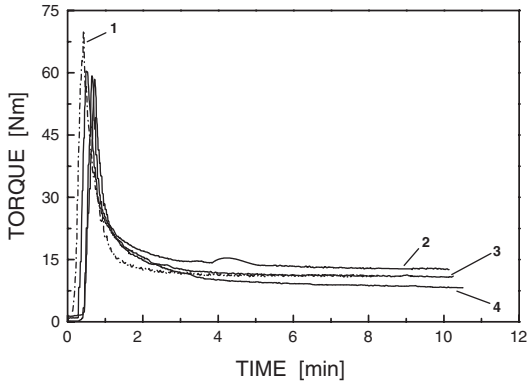


Figure 27. The torque–time curves of (1) PP (MI = 8.5 g/10 min), (2) SiO_2-g-PS/PP, (3) SiO_2-g-PMMA/PP, and (4) SiO_2 as-received/PP. SiO_2 content = 3.31 vol.%. SiO_2-g-PS and SiO_2-g-PMMA denote SiO_2 grafted by polystyrene and polymethyl methacrylate, respectively. Reprinted with permission from [107], M. Z. Rong et al., *Polymer* 42, 167 (2001). © 2001, Elsevier Science.

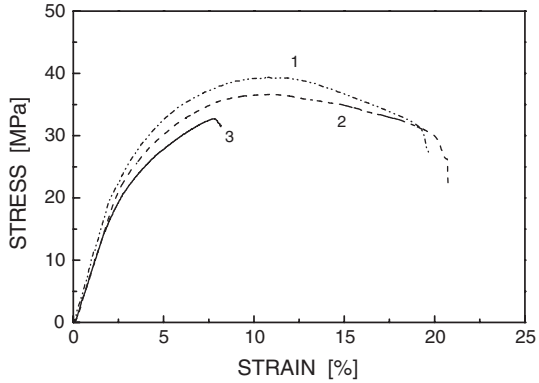


Figure 28. Typical tensile stress–strain curves of PP (MI = 6.7 g/10 min) composites filled with (1) SiO_2-g-PS, (2) $CaCO_3$-g-PS, and (3) neat PP. Content of SiO_2 or $CaCO_3$ = 1.96 vol.%. $CaCO_3$-g-PS denotes $CaCO_3$ grafted by polystyrene. Reprinted with permission from [153], M. Z. Rong et al., *J. Mater. Sci. Lett.* 19, 1159 (2000). © 2000, Kluwer Academic.

Some researchers added nanoparticles without any surface treatment into polymers. For example, Schwartz and Bahadur directly blended γ-Al_2O_3 (33 nm) with polyphenylene sulfide (PPS) powder in an electric mixer for 60 s. The mixture was then compression molded in a die at 310 °C [154]. Flexural tests on the composites for 1, 2, 3, and 5 vol.% alumina nanoparticle proportions indicated that their flexural strengths were lower than that of the neat PPS. The strength decrease was attributed to the weakening inevitably introduced by the presence of the heterogeneous particles. However, untreated nanoparticles are not always bound to result in a negative influence. Sumita et al. used a two-roll mill to mix silica with different diameters (7, 16, and 40 nm, respectively) and PP at 180 °C for 15 min [155]. The mixture was then molded at 200 °C under a pressure of 10 MPa. It was demonstrated that 20 wt.% of the nanoparticles increased the tensile yield strength of PP by 55–75% and its Young's modulus by 35–69%, while micrometer-sized inclusions only marginally increased the yield stress and the modulus. In addition, the strength and Young's modulus increased with increasing nanofiller content and decreasing

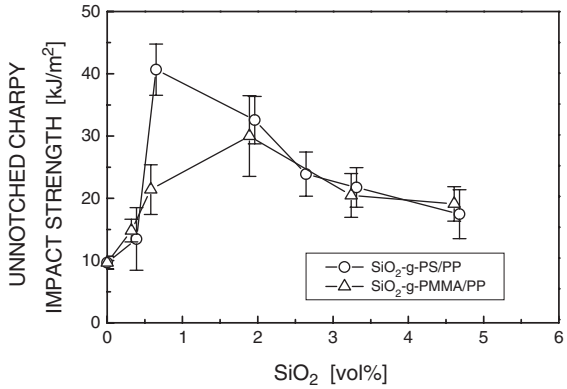


Figure 29. Filler content dependence of unnotched Charpy impact strength of PP (MI = 6.7 g/10 min)-based nanocomposites filled with various nano-SiO_2. Reprinted with permission from [107], M. Z. Rong et al., *Polymer* 42, 167 (2001). © 2001, Elsevier Science.

nanofiller size. They observed similar effects of the same nanoparticles in nylon 6 [156], PMMA [157], and PVC [158]. In the works of the authors, focused on the elastic performance of nano-SiO_2-filled PP [159] and LDPE [40], the above reinforcing behavior proved to be valid for different conditions. By comparing a modified expression, based on the dispersion strength theory, with the measured filler loading dependence of the yield stress, Sumita et al. were able to determine an aggregation parameter. It was thus concluded that the extent of nanoparticle aggregation did not increase in a polar matrix within a range of the filler volume fraction between 0 and 0.1. In a nonpolar matrix, however, the aggregation became larger with an increasing amount of particles. But this analysis has to be taken cautiously, since no direct microstructural investigation was conducted.

Landry et al. also mixed untreated fumed silica (7 nm) and molten polyvinyl acetate (PVAc) at about 100 °C for 20 min in a two-roll mill [160]. The mixture was melt-pressed at 100–120 °C to obtain thin, optically transparent composite films. At a filler loading of 18 wt.%, room-temperature tensile modulus and tensile stress of the composites were increased by 30% and 5%, respectively, but the elongation to break was reduced by a factor of 3.5, when compared with the neat PVAc control. The marginal improvement of the stiffness and strength was interpreted as a lack of a "percolated" SiO_2 network.

With an improved understanding of carbon nanotube properties and the continuous development of nanotube production and purification, there is an increasing in fabricating composites containing carbon nanotubes [161]. Cooper et al. assumed that the difficulty of producing well-dispersed carbon nanotubes in a composite is due to a phase separation of the carbon and polymer powders. The latter can occur as the result of sonication of a small amount of solid (carbon) in a comparatively large amount of liquid. The carbon nanotubes are more attracted to each other than they are to the larger volume of the liquid. To solve the problem, Cooper et al. suggested that the two solid phases of the composite should be combined in a relatively dry state [162, 163]. In this "dry" processing method, even though a very small amount of liquid is used to transfer sonication energy to the particles (carbon and matrix), there is only a low likelihood that the nanotubes will pass around the solid matrix particles to agglomerate together. Based on the above idea, multiwalled carbon nanotubes (MWNTs) and single-walled carbon nanotubes (SWNTs) were mixed by ultrasonication in a small amount of ethanol to achieve a homogeneous slurry of spherical PMMA particles (100–500 μm). The mixture was dried and mechanically mixed with an attritor operating at 2000 rpm. This mixing was followed by a kneading process at 170 °C, so that the matrix became a viscous mass. The final step of composite preparation was to extrude the premixture to orient the reinforcement and the polymer matrix in the flow direction. Tensile tests indicated that the Young's modulus of the MWNT composites was even lower than that of the neat PMMA. This was attributed to the fact that the inner nanotube shells were sliding past each other, thus preventing effective load transfer. This was confirmed by Raman spectroscopy, since no shift in the peak position of the A_{1g} Raman band for MWNTs was observed when a nanotube composite was deformed in tension [164]. Nevertheless, the impact resistance of the SWNT composites at a rather low filler content of 0.01 wt.% was about four times higher than the value of the matrix. It is believed that the high flexibility and elasticity of the nanotubes increase the fracture toughness of the composite by absorbing the impact energy.

Table 9 summarizes the polymer nanocomposites discussed in this section in terms of component species, filler treatment, and ultimate properties of the composites. It is seen that (a) $CaCO_3$ nanoparticles are the most frequently used fillers, and (b) a coupling agent treatment is the main way to enhance the interfacial interaction between the nanoparticles and the matrix polymers. This statistic seems to be similar to the classic situation of microscale particulate composites. Many techniques for modifying nanoparticles, as intensively reviewed previously, have not yet been applied to practical manufacture. In fact, incorporation of nanoparticles into polymers has scarcely brought about remarkable improvement in overall properties of the matrices in most systems investigated so far. Experience has shown that the reproducibility of performance enhancement is not always guaranteed. This means that mass production of composites is not yet available. Continuous efforts must be made to develop new techniques to bring the extraordinary characteristics of nanoparticles into full play in the plastics industry. Mechanically copying the common practice in the manufacture of microparticle-filled polymer composites would not result in a genuine polymer nanocomposite [165].

3.2.2. Solution Mixing

Solution mixing is a liquid-state powder processing method (particles in a polymer solution) that brings about a good molecular level of mixing, but at a cost depending on the solvent and its recovery. It is applicable to the polymers that can be dissolved or swelled by the solvent. Its advantages include large interface extension, improved interfacial bonding, low viscosity, fast solvent removal, easy casting, and high homogeneity of the ultimate composite materials.

Carotenuto et al. produced monolithic silica (356 nm)/PMMA nanocomposites through solution blending [166]. They mixed a silica/methoxy propylacetate suspension with a PMMA solution (50 wt.% polymer in methoxy propylacetate) and then sonicated the mixture for 5 min. The samples were centrifuged to remove bubbles and then dispersed into a mold. After being heated in a vacuum at 100 °C for 15 min, the nanocomposites were ready for use. The particle surface modification was obtained by treatment of the silica particles with an organic colorant (methyl red) at room temperature [167]. Electron microscopy indicated the formation of a polymer coating around the organically modified particles in the starting liquid precursor, and the resultant composites exhibited an excellent homogeneity. In contrast, the untreated SiO_2 formed a lot of agglomerates ($\sim$1 μm) in the matrix after evaporation of the solvent due to the weak hydrogen bonds between the methyl-ester groups of the polymer and the hydroxyl groups on the particle surface. A significant improvement of the Young's modulus was observed after the introduction of the nanofiller. The Vickers hardness of a composite with 50 wt.% silica was increased by 90%, and the glass transition temperature

Table 9. Variation in mechanical performance of some polymer nanocomposites manufactured through melt mixing as compared with unfilled polymers.

Nanosized fillers	Filler pretreatments	Polymer matrices	Property variation[a]							Ref.
			Young's modulus	Flexural modulus	Tensile strength	Flexural strength	Elongation to break	Impact strength	K_{IC} and/or G_{IC}	
$CaCO_3$ (70 nm)	Stearic acid coated	PP	+	N/A	−	N/A	N/A	N/A	+	[142]
$CaCO_3$ (70 nm)	Titanate coupling agent	PP	+	N/A	−	N/A	N/A	N/A	+	[142]
$CaCO_3$ (50 nm)	Titanate coupling agent	HDPE	N/A	N/A	−	N/A	+	+	N/A	[136, 137]
$CaCO_3$ (50 nm)	Macromolecular coupling agent	HDPE	N/A	N/A	−	N/A	++	+	N/A	[137]
$CaCO_3$ (80 nm)	Phosphonate coupling agent	PP	+	+	+	N/A	N/A	++	N/A	[140]
$CaCO_3$ (30 nm)	Aluminate coupling agent	PVC	N/A	N/A	+	N/A	−	+	N/A	[165]
$CaCO_3$ (30 nm)	Aluminate coupling agent	Acrylate resin/PVC	N/A	N/A	++	N/A	+	+	N/A	[165]
SiO_2 (7 nm), $CaCO_3$ (60 nm)	Irradiation grafting	PP	+	N/A	+	N/A	+	++	++	[107, 153]
SiO_2 (20 nm)	Silane coupling agent	PMMA	N/A	N/A	−	+	−−	+	N/A	[138]
SiO_2 (7, 16, 40 nm)	Untreated	PP, LDPE, PVC, Nylon 6, PMMA	+	N/A	+	N/A	N/A	N/A	N/A	[40, 155–159]
SiO_2 (7 nm)	Untreated	PVAc	+	N/A	+	N/A	−−	N/A	N/A	[160]
SiC/Si_3N_4(20 nm)	Titanate coupling agent	LDPE	N/A	N/A	++	N/A	+	++	N/A	[134]
SiC/Si_3N_4(20 nm)	Titanate coupling agent	CPE/PVC	N/A	N/A	+	N/A	N/A	++	N/A	[135]
TiO_2 (20 nm)	Macromolecular dispersant	HIPS	+	N/A	+	N/A	+	+	N/A	[150]
Al_2O_3 (33 nm)	Untreated	PPS	N/A	N/A	N/A	−	N/A	N/A	N/A	[154]
Carbon nanotubes	Untreated	PMMA	−[b]	N/A	N/A	N/A	N/A	++[c]	N/A	[162, 163]

[a] The symbol + corresponds to a moderate rise in the properties, ++ to a prominent improvement, and − describes a moderate decrease in the properties and −− to a prominent decay.
[b] Multiwalled carbon nanotubes serving as the fillers.
[c] Single-walled carbon nanotubes serving as the fillers.

became 5 °C higher than the characteristic value of the pure PMMA.

With the same mixing technique, Carotenuto et al. prepared nanocomposites consisting of poly(methacrylic methylester) and monodispersed $Cu_2(OH)_2CO_3$ particles [168]. They found that the thickness of the polymeric layer around the particles increased with the polymer concentration. Hence it may be possible to manipulate the composite morphology and properties *in-situ* by changing the degree of wrapping of the organic molecules on the inorganic filler particles.

Because of the high mobility of nanoparticles in a polymer solution, composites with specific spatial distribution of the nanoparticles can thus be made, based on the self-organization process. Hamdoun et al. prepared stable solutions of polystyrene-polybutylmethacrylate P(S-b-PBMA) diblock copolymer (M_W = 82,000, molecular weight polydispersity index = 1.05) and γ-Fe_2O_3 (3.5 nm) in toluene [169]. To acquire good solubility, the nanoparticles were precovered by a PS layer of short chains (M_W = 13,000). Eventually, a thermodynamically stable nanocomposite was obtained in three steps: (i) mixing of the two solutions; (ii) spin-coating the composite solution on a solid substrate; (iii) annealing the resulting film at 150 °C under vacuum, until a quasi-equilibrium self-organization was reached. Microscopic observation revealed a periodic arrangement of the particles in the polymer matrix, which reproduces the ordered arrangement of the copolymer mesophase. It is expected that the macroscopic properties of the composite could reflect both the physical properties characteristic of the nano-objects and those that are specific to the large mesh periodic structure of the matrix.

During a series of works by Shang et al. focusing on the effect of interfacial bonding on the mechanical performance of particulate filled polymer composites [41, 170, 171], a silica/ethylene vinyl acetate (EVA) composite system was manufactured by means of a solution mixing technique. The silica particles, either 600 nm or 14 nm in diameter, were modified by a thermal treatment at different temperatures or by a thermal treatment combined with a chemical treatment with trimethylchlorosilane. Then the particles were incorporated into the EVA/benzene solution. For the coarser silica composite solutions, a magnetic stirrer was used to improve the particle dispersion. For the finer particles, sonication was employed instead. Benzene was chosen as a solvent because it is a good and nonpolar solvent for EVA copolymer. The silica particles were expected to be well distributed in the polymer solution since the solvent

induces optimal polymer chain length expansion that stabilizes the silica powders in solution [171]. The tensile strength values of the composites indicated that the SiO_2 (600 nm)-filled versions had a lower tensile strength than the unfilled EVA, suggesting that the particles had a weakening effect. In contrast, the composites filled with SiO_2 (14 nm) were stronger than the matrix polymer. Within the lower loading range of SiO_2 (14 nm), the tensile strength of the composites increased with increasing volume fraction of filler, reaching a maximum at around 4 vol.%. The composites retained their higher tensile strength compared with the unfilled polymer up to 15 vol.% of filler [41]. On the other hand, the Young's moduli of the composites increased with a rise in the content of either SiO_2 (600 nm) or SiO_2 (14 nm) [170]. For identical filler volumes, the SiO_2 (14 nm) composites were stronger than the SiO_2 (600 nm) ones, owing to the fact that the fine particles have more surface area available to bond and adsorb matrix polymer. These bonded and adsorbed polymer chains become stiff as a result of a loss in flexibility.

To diminish mixing problems and to improve the dispersion of fine particles in polymers, Vollenberg and Heikens combined solution mixing and melt mixing when preparing alumina-filled composites [172]. They first made master batches containing 30 vol.% of alumina beads (35 nm and 400 nm in diameter, respectively) by adding the particles to the polymer solutions. Having been stirred for a few hours, the mixture was poured onto a large surface, allowing the evaporation of the solvent overnight. The master batches were then dried at 100 °C under vacuum. Finally, the unfilled polymers were mixed with the master batches on a two-roll mill to obtain the desired composites. The polymers employed in their work were PS, styrene-acrylonitrile copolymer (SAN), and polycarbonate (PC). The solvent for the former two polymers was ethylene acetate, and methylene chloride was used to dissolve the latter one. Similar to the case stated above, the Young's moduli of the composites appeared to be dependent on the particle size. At a constant volume fraction of filler, the smaller alumina particles (35 nm) generated a more significant reinforcing effect than the larger ones (400 nm).

To compare the properties of nanocomposites, prepared through a sol–gel route, with particulate-filled systems, Landry et al. added nano-SiO_2 (7 and 14 nm, respectively) to a 10 wt.% solution of PVAc in tetrahydrofuran (THF) and mixed them for several hours [160]. The solution was then cast in a Teflon mold, yielding opaque, broken pieces of the composites. Eventually, transparent nanocomposite films were obtained by melt-pressing the substance at 100–120°C. Tensile testing indicated that below the glass temperature, T_g, of PVAc, the solution-mixed samples possess properties similar to those of the melt-mixed ones (see last section). At a temperature higher than T_g, however, the solution-mixed composites decayed rapidly, in sharp contrast to the rubber-like plateau of the *in-situ* composites extending to about 300 °C. Again these observed differences were attributed to the differences in particle–particle connectivity.

Shaffer and Windle fabricated carbon nanotube composite films by carefully mixing aqueous poly(vinyl alcohol) (PVA) solutions with catalytically grown carbon nanotube dispersions followed by subsequent casting and controlled water evaporation [173]. The electrostatically stabilized dispersion of carbon nanotubes in water was produced when the nanotubes had been pretreated by a suitable oxidation [174]. Successful preparation depended crucially on maintaining stable colloidal mixtures of nanotubes and polymer. Dynamic mechanical analysis of the composites suggested that the presence of carbon nanotubes stiffens the material, particularly at high temperature, and in some cases retards the onset of thermal degradation. The authors concluded that carbon nanotubes might not ideally be suited to a straightforward reinforcing role, but that they could find application as a modifier for polymers, particularly as an improved matrix for conventional fiber composites for service at high temperatures.

3.2.3. Miscellaneous Mixing Techniques

Thermal spraying seems to be a viable solution to the processing limitations of conventional techniques for polymer nanocomposite manufacturing, such as the use of a high working temperature for melt mixing and solvents for solution mixing [175]. This is particularly true for the preparation of nanocomposite coatings. During a thermal spray process, materials, in powder, wire, or rod form, are heated, accelerated, and propelled by a high-temperature jet through a confining nozzle toward a surface. The individual molten or softened droplets impact, spread, cool, and solidify to form continuous coatings. Material heating, quenching, consolidation, and posttreatment are thus combined in a single step. The important benefits of using a thermal spray process for making polymer nanocomposites include the following: (i) powdered polymers are heated only to a temperature at which the particles are viscous enough to spread because of the high kinetic energy; (ii) good dispersion of nanoparticles is achieved by the high velocity; (iii) processing without the use of a volatile solvent is possible.

Schadler et al. employed a high-velocity oxy-fuel (HVOF) system to thermally spray SiO_2/nylon 11 nanocomposite coatings [175]. The particles (7 and 12 nm in diameter, respectively) were pretreated to be either hydrophobic or hydrophilic, respectively. To avoid powder segregation in the jet due to differences in powder size and density, the nylon and silica particles were dry ball milled together for 48 h, resulting in a gradient shell of embedded silica particles near the surface of the nylon powder particles. By optimizing spray parameters such as nozzle design, spray distance, oxygen-to-fuel ratio, powder feed position, and substrate cooling, dense coatings with relatively uniform particulate distribution were achieved. Compared with neat nylon, scratch resistance improved by 30% and wear resistance improved by 55%. Silica particles with a hydrophobic surface led to better mechanical properties than those with a hydrophilic surface.

Recently Petrovicova et al. characterized the changes that the above silica reinforced nylon 11 underwent during thermal spraying and their influence on the final polymer morphology and coating performance [176, 177]. Agglomerates of silanated silica, on the order of 50 nm in size, were observed in the nanocomposites, whereas the agglomerates of untreated silica and hydrophilic silica were on the order of 100 nm. Reinforcement of the polymer matrix, for

example, in the case of a 15 vol.% silanated silica-filled nanocomposite, resulted in an increase of up to 205% in the dynamic storage modulus (Table 10). At temperatures above the glass transition temperature of nylon 11, the storage modulus of a 15 vol.% hydrophobic silica nanocomposite was higher than that of the unfilled version by up to 195%. An increase in matrix crystallinity was assumed to be the predominant cause of the enhancement of the mechanical properties.

Unlike conventional dispersive mixing at a temperature above the melting point or the flow temperature of the polymer matrix, mechanical alloying, which compounds the components at room temperature through mechanochemical effects, is an effective route for producing polymer composites with sufficiently high interfacial adhesion. The process is a high-energy ball-milling one in which the repeated fracture and welding of powder particles, arising from ball-powder collision events, allows true alloy powders to be formed from mixtures of elemental powders [178, 179]. It appears to be very promising because it offers the possibility of mixing two ordinarily immiscible materials, and no coupling agents or other surface treatments are needed, in principle.

So far, very few results have been reported that deal with mechanical alloying of polymer nanocomposites. Hu et al. applied mechanical milling to the preparation of composites consisting of carbon nanotubes (20 nm in diameter) and an ultra-high-molecular-weight polyethylene (UHMWPE) [180]. Having been milled for 2 h, the mixture of carbon nanotubes and the micron powder of UHMWPE was homogeneous. At a filler loading of 1 wt.%, the impact strength of the composites increased by a factor of 43% compared with the value of the unfilled UHMWPE. Electron microscopic observations of the fractured surface indicated that a resin layer (15–20 nm thick) was wrapped around the nanotubes, suggesting a good adhesion between the fillers and the matrix formed during mixing.

Table 10. Dynamic storage modulus G' of thermally sprayed nanocomposite coatings.

Silica content (vol.%)	Silica surface	Polymer matrix	G' at 30 °C (GPa)	G' at 70 °C (GPa)
0	—	D-60[a]	0.65	0.36
0	—	D-30[b]	0.86	0.35
5	Hydrophobic	D-60	1.15	0.86
5	Hydrophilic	D-60	1.58	1.07
5	Silanated[c]	D-60	1.25	0.39
10	Hydrophobic	D-60	1.10	0.89
10	Hydrophobic	D-30	1.38	0.59
10	Hydrophilic	D-30	1.38	0.54
10	Silanated	D-60	1.29	0.61
15	Hydrophobic	D-60	1.39	1.06
15	Hydrophobic	D-30	1.83	0.61
15	Hydrophilic	D-60	1.31	0.99
15	Hydrophilic	D-30	1.79	0.56
15	Silanated	D-60	1.98	0.83

[a] D-60 denotes nylon 11 powders with a mean particle size of 60 μm.
[b] D-30 denotes nylon 11 powders with a mean particle size of 30 μm.
[c] Silanated silica means the silica surface was modified with the use of A-1100 γ-aminopropyltriethoxy silane.
Source: Reprinted with permission from [177], E. Petrovicova et al., *J. Appl. Polym. Sci.* 78, 2272 (2000). ©2000, Wiley-VCH.

Chen et al. found that α-Fe_2O_3 nanoparticles (10 nm) were generated in the ball milling process of a Fe_3O_4/PVC system in the presence of air [181]. With the use of Mößbauer spectroscopy, it was revealed that PVC powders were partially degraded because of mechanical milling and became reactive on their surface. α-Fe_2O_3 nanoparticles were the products of activated PVC and Fe_3O_4. The authors also found that Mößbauer scattering, characteristic of interfacial interactions between Fe_3O_4 and PVC, appeared even in the case of argon protection.

4. MECHANISMS FOR PROPERTY IMPROVEMENTS

4.1. Interfacial Interaction

An extremely high surface area is one of the most attractive characteristics of nanoparticles because it facilitates the creation of a great amount of interphase in a composite. Introduction of nanoparticles into a polymer changes the intermolecular interaction of the matrix [182]. As estimated by Reynaud et al. [5], an interphase, 1 nm in thickness, represents roughly 0.3% of the total volume of the polymer in the case of microparticle-filled composites, whereas it can reach 30% of the total volume in the case of nanocomposites. Therefore, the nonnegligible contribution made by the interphase provides diverse possibilities for performance tailoring.

To study the influence of the interphase on the properties of polymer composites, two groups of models have been proposed. The first one assumes uniform properties of the interphase [183–185], and the other assumes nonuniform properties [186, 187]. It seems that the nonuniform models of the interphase are limited to relatively large particles. For nanoparticles, the interphase region is on the order of some nanometers [188]. In this context, it is reasonable to assume that the physical properties of the interphase in nanocomposites are either uniform or nearly uniform [189].

The nature of interfacial interactions is a function of atomic arrangement, molecular conformation, chemical composition, morphology, and diffusivity of the filler and the matrix [190]. Usually the strength of the interaction is created by primary bonds (e.g., ionic, covalent, and metallic) and/or secondary bonds (e.g., hydrogen bonds and van der Waals forces) [54]. With respect to the mechanisms responsible for interfacial adhesion in microcomposites, numerous theories have been studied, including adsorption and wetting, electrostatic attraction, chemical bonding, interdiffusion, mechanical interlocking, etc. However, only a few papers have discussed the interfacial effects of nanocomposites because the importance of the materials began to be recognized just in recent years.

According to the theory of interdiffusion, two surfaces in contact with each other can build up adhesion by the mutual diffusion of the molecules of the interacting surfaces. The bond strength is dependent on the amount of molecular entanglement, the number of molecules involved, and the strength of the bonding between the molecules. The interphase thus formed has a substantial thickness, and its chemical, physical, and mechanical properties are different from those of the filler and the matrix. Kendall and

Sherliker demonstrated for nanosilica/polyethylene composites the presence of "bound polymer," which cannot be removed by solvent extraction from the filled material [188]. It was suggested that there are two interfaces; one is between the filler and the bound polymer, and a second interface occurs between the bound polymer and the soluble polymer. The latter acted as a weak point, embrittling the matrix through the reduced entanglement resulting from the surface assymmetry. To overcome this problem, increasing the molecular weight of the matrix polymer to produce a better physical entanglement across the second interface proved to be effective [188, 191].

In fact, the expectation that a sufficiently long surface chain, extending some distance from the filler, would strengthen the interfacial bonding [191] was further confirmed by Rong et al. for nanoparticle-filled PP [107]. They introduced macromolecules with a molecular weight of around 10^4 onto nano-SiO_2 (7 nm) through irradiation grafting. Although the monomers of the grafting polymers should have different miscibilities with PP, all of the grafting polymers, except PEA, exhibited a reinforcement effect on the tensile strength of the nanocomposites (Table 11). These results contribute to a further understanding of the modified nanoparticles and their role in the composites. That is, interdiffusion instead of a miscibility of the grafting polymer segments with the polypropylene molecules, and entanglement between the grafting polymer and the matrix, dominate the interfacial interaction in the nanocomposites. This leads to the conclusion that a PP matrix with a higher molecular weight should entangle even more effectively with the nanoparticles agglomerates, thus leading to a higher tensile strength increment. The experimental data of the authors for PP with different molecular weights support this estimation.

In a recent report on grafted nano-SiO_2 particle-filled PP composites studied by atomic force microscopy (AFM) [192], it was shown that PP molecules are able to immigrate into the nanoparticle agglomerates in the molten state of PP (Fig. 30). An interdiffusion and entanglement between the molecules of the grafting polymers and the matrix can thus take place outside and inside the nanoparticle agglomerates. Moreover, Shumsky et al. [193] measured the viscoelastic behavior of the nanocomposites employed in [107] in the molten state. Drastic changes in the pattern and broadening of the relaxation times spectra to longer relaxation times for the composites were attributed to the onset of plastic yielding of a spatial network of the filler particles grafted with macromolecular chains, preceding the macroscopic melt flow. The results of this study can be regarded as experimental evidence for a shear-resistant, stiff spatial structure of filler particles coated with a polymer boundary layer in the melt phase of a nanocomposite with relatively low filler content. Landry et al. also proposed the existence of a nanosilica network in PVAc-based composites [160]. The chains of PVAc are strongly adsorbed to and entrapped in the silicate network, as suggested by a high plateau modulus above T_g, a broadened distribution of relaxation times at T_g, a diffuse interface between silica and PVAc, and a high level of hydrogen bonding.

Kunz and co-workers studied the distributions of colloidal gold nanoparticles (20 nm) in poly(2-vinylpyridine) (PVP), at the surfaces of PS and PVP, and at the interface between PS and PVP [194]. Contrary to the results obtained for PS surfaces, it was found that the particles adsorb strongly to the PVP surface, even at temperatures at which gold particles are quite mobile. A striking manifestation of slow dynamics at the gold/polymer interface is the pinning of gold nanoparticles to an interface between PS and PVP. The effect can be explained by the formation of loops between different portions of an effectively grafted PVP chain. PS chains are linked to the gold particles by the entanglements they make with these loops.

Espiard et al. further described the details of molecular entanglements at the interfacial region of a nano-SiO_2 reinforced poly(ethyl acrylate) (PEA) film to explain the improved mechanical properties [195]. The nanoparticles used in their work were pretreated with silane carrying a polymerizable methacrylic group, which furthermore served as a nucleating agent in an emulsion polymerization of ethyl acrylate monomers [80, 196]. It was found that grafting of ethyl acrylate to the treated nano-SiO_2 particles took place in the earlier stage of the polymerization. As a result, the polymer that had not been grafted participated in two kinds of entanglements. One involved polymer molecules coming from different particles. Entanglements are formed during the coalescence process of the film formation. The other involved both the nongrafted polymer and the tight loop system formed around the particles by copolymerization of the methacrylic groups of the coupling agent and the growing polymer. Upon elongation of the film, the set of entanglements involving the loops is expected to remain, whereas the nongrafted polymer can be elongated extensively and

Table 11. Mechanical properties of PP (MI = 8.5 g/10 min)-based nanocomposites[a] filled with SiO_2 grafted with different polymers.

	Nanocomposites						Neat
Grafting polymers	PS	PBA	PVA	PEA	PMMA	PMA	PP
Tensile strength (MPa)	34.1	33.3	33.0	26.8	35.2	33.9	32.0
Young's modulus (GPa)	0.92	0.86	0.81	0.88	0.89	0.85	0.75
Elongation to break (%)	9.3	12.6	10.0	4.6	12.0	11.9	11.7
Area under tensile stress–strain curve (MPa)	2.4	3.3	2.3	0.8	3.2	2.9	2.2
Unnotched Charpy impact strength (kJ/m^2)	19.8	19.4	22.9	14.6	20.5	4.7	8.0

[a] Content of SiO_2 = 3.31 vol.%.

Source: Reprinted with permission from [107], M. Z. Rong et al., *Polymer* 42, 167 (2001). © 2001, Elsevier Science.

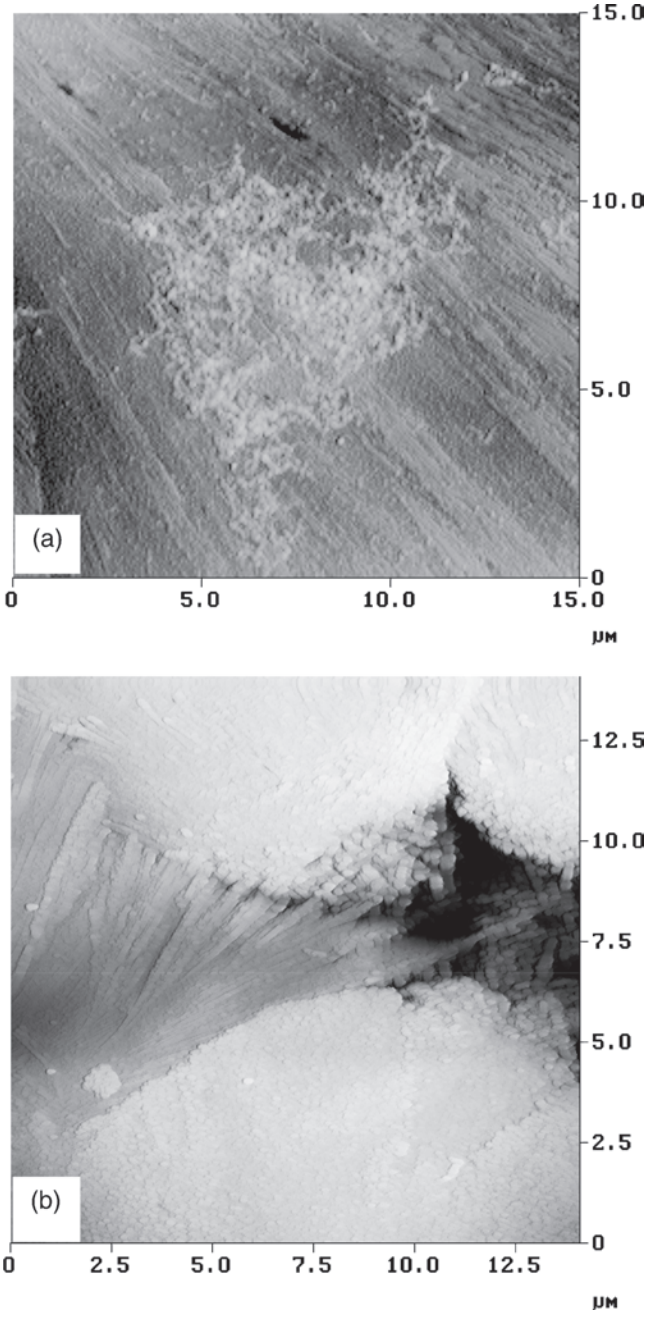


Figure 30. AFM images of etched SiO_2-g-PS/PP. (a) Agglomerated SiO_2-g-PS dispersed in the lamellae of PP. (b) PP lamellae penetrating the cave resulting from the etching of SiO_2-g-PS. The remaining nanoparticles at the edge of the cave indicate that some nanoparticle agglomerates have been there. Reprinted with permission from [192], M. Q. Zhang et al., *J. Appl. Polym. Sci.* 80, 2218 (2001). © 2001, Wiley-VCH.

quasi-reversibly up to near the rupture process, which occurs when the first set of entanglements disappears.

The above-cited works imply that the interfacial interaction in nanocomposites is often a combination of different effects, like adsorption, chemical bonding, interdiffusion, etc. It is hard to treat the problem based on a single mechanism. Shang et al. proved the existence of intermolecular forces between the silanol groups of the nanosilica filler and the carbonyl groups of the EVA matrix [41]. A reduction in the amount of surface hydroxyl groups of nanosilica results in a weaker interfacial bonding. Within the nanocomposites of SiO_2/EVA, more EVA molecules are adsorbed and coupled through entanglements other than direct hydrogen bonds to the silica surface [170]. In the course of rheological measurements at 150 °C, the adsorbed polymer gradually loses its entanglements through decoupling at the silica surface with a rise in the testing frequency [171].

Since many of the thermoplastics, serving as matrices in nanocomposites, are semicrystalline (with properties that are strongly related to the crystalline characteristics), effects of a heterogeneous nucleation exerted by the nanoparticles would significantly influence the microstructure of the matrix polymers. This in turn affects the mechanical performance of the nanocomposites. In a work on nanosized carbon fiber reinforced polyether polyurethane, Xu and co-workers detected two melting peaks of the matrix polymer in the composites, instead of a single melting peak of the polymer without a nanoparticle reinforcement [197]. Ren et al. measured the melting behavior of a $CaCO_3$ (80 nm)/PP nanocomposite in comparison with that of the neat PP [149]. They found a minor melting peak corresponding to the β-phase PP at 150.6 °C in addition to the main endotherm at 165 °C typical for the α-phase PP (Fig. 31). Since impact strength and toughness of β-PP exceed those of α-PP because of β-to-α or a β-to-smectic transition during the deformation [198, 199], an appearance of β-phase in PP should account for the improvement in ductility of the nanocomposites.

In general, γ-quinacridone is the most widespread highly active β-nucleating agent. Some two-component compounds obtained by the reactions of organic acids with $CaCO_3$ can also serve as β-nucleators [200, 201]. Moreover, Liu et al. indicated that certain inorganic nanoparticles like Y_2O_3 (20 nm) are able to induce β-crystallites in PP [202]. At an Y_2O_3 content of 1 wt.%, the melting enthalpy of β-PP approaches 3/4 compared with that of α-PP [203]. Nevertheless, the role of nanoparticles in modifying the crystallization habit remains unclear because research groups

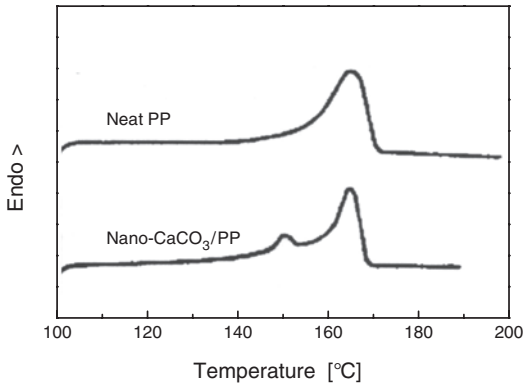


Figure 31. Differential scanning calorimetry (DSC) heating traces of PP and its nanocomposites. Both specimens were prepared at 130 °C. Reprinted with permission from [149], X. Ren et al., *Chin. Plast.* 14, 22 (2000) (in Chinese). © 2000, Chinese Plastics Processing Industry Association.

have obtained different results. Wu and Xu examined the crystallization behavior of SiO_2 (20 nm)/PP/ethylene-propylene-diene monomer (EPDM) terpolymer [204]. Nano-SiO_2 was found to increase the crystallization temperature and crystallization rate of PP and to reduce the spherulite size. However, there was no trace of a β-phase PP. Similarly, Rong et al. demonstrated that both the untreated and treated nano-SiO_2 have some nucleation influence on the crystallinity of PP [107]. Except for this effect, the difference between the nanocomposites and the neat PP could not be detected during the nonisothermal and isothermal crystallization measurements (Tables 12 and 13). Therefore, the reinforcing and toughening effects of the modified nanoparticles could not be attributed to the slight variation in matrix microstructure.

4.2. Quantitative Considerations

Although there are quite a lot of theories for the mechanical properties of composite materials, usually they do not offer a general solution. In the case of nanocomposites, for example, in which the nanofillers are used to modify the properties of the polymer matrix at a rather low filler concentration, it is sometimes rather difficult to describe the performance by using the theories for conventional polymer composites [205]. Specified models have to be worked out as a result. Nevertheless, nanocomposites are a rather new member of the family of composite materials, and for the moment continuous efforts are needed to collect as much information as possible about their deformation behavior and damage processes. We are far from the stage of concluding theoretical estimations with adequate applicability and predicability. Instead, analyses of case studies would help to establish a knowledge path and a frame of methodology for dealing with a quantitative description of the mechanical role of nanoparticles in composites.

4.2.1. Stiffness

The rule of mixtures is a general approximation of the description of the mechanical behavior (especially elastic behavior) of composites. It gives upper (Eq. (6)) and lower (Eq. (7)) bounds to their mechanical response:

$$E_c = \phi_m E_m + \phi_f E_f \tag{6}$$

$$E_c = E_m E_f / (\phi_m E_m + \phi_f E_f) \tag{7}$$

where E_c is the modulus of the composites, E_m and E_f denote the moduli of the matrix and the filler, and ϕ_m and ϕ_f symbolize the volume fraction of the matrix and the filler, respectively. As a matter of fact, Eqs. (6) and (7) are based on a strong interfacial adhesion and rigid spherical reinforcement, respectively. The main reinforcing mechanism is attributed to the efficiency of the load transfer from the matrix to the filler. It was found that the moduli of most particulate filled polymers lie between the two limits [206]. On the other hand, Kerner's equation was derived from a basic consideration of mechanical properties of a composite at moderate concentration [207]. For low filler content, Nielsen's modification of Kerner's equation is more applicable because the maximum packing of the fillers, interfiller interactions, and the mechanical coupling of the matrix and filler are taken into consideration [208, 209]:

$$E_c = E_m(1 + AB\phi_f)/(1 - B\psi\phi_f) \tag{8}$$

where A accounts for the contribution of the interfacial interaction and factors like filler geometry, B represents the relative stiffness of the two components, and ψ is dependent on the filler packing fraction.

Predictions of nanocomposite moduli based on the above equations were made by Petrovicova et al. [177]. The experimental results indicated that all of the nano-SiO_2/nylon 11 nanocomposites exhibited a modulus higher than that calculated by the rule of mixtures. Although the data for most of the nanocomposites closely fit Eq. (8) for $A = 3 - 4$, there was a large deviation for the 2.4 vol.% filled composites, manifesting the extraordinary reinforcing tendency of the nanoparticles. It was believed that the lower values of A were due to an improved spatial distribution of the silica in the matrix.

As a result of the unsatisfactory agreement between the Kerner equation and the measured data, Kryszewski and Bak [189] assumed that some volume of the surrounding

Table 12. Nonisothermal crystallization and melting data of PP (MI = 8.5 g/10 min) and its nanocomposites.

Samples	Content of SiO_2 (vol.%)	T_m[a] (°C)	T_{cn}[b] (°C)	ΔT[c] (°C)	X_c[d] (%)
Neat PP	—	164.7	115.3	49.4	44.6
Nanocomposites filled with untreated SiO_2	1.96	165.8	117.8	48.0	46.4
	3.31	166.6	118.9	47.7	45.1
	4.68	164.4	118.6	45.8	44.4
	6.38	163.3	118.9	44.4	46.7
Nanocomposites filled with SiO_2-g-PS	1.96	164.7	116.5	48.2	45.8
	3.31	166.2	118.9	47.3	43.9
	4.68	164.9	118.7	46.2	46.9
	6.38	165.3	119.1	45.3	44.6

[a] T_m denotes the peak melting temperature.
[b] T_{cn} denotes the peak crystallization temperature recorded during cooling.
[c] $\Delta T = T_m - T_c$, denotes the supercooled temperature.
[d] X_c denotes the crystallinity of PP.
Source: Reprinted with permission from [107], M. Z. Rong et al., *Polymer* 42, 167 (2001). © 2001, Elsevier Science.

Table 13. Kinetic parameters of isothermal crystallization of neat PP (MI = 8.5 g/10 min) and its nanocomposites.

Samples	T_{ci}[a] (°C)	ΔH[b] (J/g)	t_f[c] (min)	t_{max}[d] (min)	$t_{1/2}$[e] (min)	n[f]	K[g] (min^{-n})	T_m (°C)
Neat PP	130	95.6	14.6	5.7	6.1	2.69	5.30×10^{-3}	166.5
	132	97.9	23.8	9.8	10.1	2.75	1.18×10^{-3}	167.1
SiO_2 as-received/PP[h]	130	92.3	10.2	4.1	4.3	3.10	7.16×10^{-3}	166.8
	132	94.7	17.1	7.5	7.9	2.94	1.64×10^{-3}	167.3
SiO_2-g-PS/PP[h]	130	93.4	11.3	4.9	5.0	2.84	7.15×10^{-3}	168.6
	132	95.2	18.7	8.9	8.8	2.99	1.07×10^{-3}	167.2
SiO_2-g-PMMA/PP[i]	130	94.6	12.4	5.1	5.2	3.02	5.87×10^{-3}	167.0
	132	94.0	18.2	7.7	8.1	2.80	1.98×10^{-3}	168.0

[a] T_{ci} denotes the present isothermal crystallization temperature.
[b] ΔH denotes the enthalpy of crystallization.
[c] t_f denotes the time at which the crystallization is completed.
[d] t_{max} denotes the time at which the crystallization rate is the maximum.
[e] $t_{1/2}$ denotes the time at which the crystallization is carried out for a half.
[f] n denotes the Avrami index.
[g] K denotes the rate constant of crystallization.
[h] Content of SiO_2 = 3.31 vol.%.
[i] Content of SiO_2 = 1.96 vol.%.
Source: Reprinted with permission from [107], M. Z. Rong et al., *Polymer* 42, 167 (2001). © 2001, Elsevier Science.

matrix becomes immobilized as a result of interfacial binding, giving rise to an effective increase in the size of the nanoparticles. The effective volume fraction of the partially immobilized material ϕ_e is given by [210]

$$\phi_e = \phi_f(1 + \Delta R/R)^3 \tag{9}$$

where R stands for the radius of nanoparticles and ΔR is the increment of the effective particle size. By replacing ϕ_f with ϕ_e in Kerner's equation for a Poisson ratio of 0.5, the elastic modulus of nanocomposites can be described as

$$E_c = [1 + 2.5\phi_e/(1 - \phi_e)]E_m \tag{10}$$

A comparison between the testing data of nano-palladium (1–2 nm)/PMMA and Eq. (10) indicated a good agreement, provided $\Delta R = 2.25$ nm. This suggests that the effective radius of the immobilized matrix environment could be much greater than that of the nanoparticles. The stronger the interfacial interaction is, the thicker is the interlayer [171]. Similarly, Sumita et al. studied the energy dissipation of PP-based nanocomposites under the circumstances of dynamic mechanical testing [159]. According to the concept of the effective volume fraction of the dispersed phase, which assumes that the "immobilized" matrix associated with the interphase does not contribute to energy loss, they yielded an expression for the relative loss modulus,

$$E_c''/E_m'' = 1/(1 - \phi_e) \tag{11}$$

where E_c'' and E_m'' represent the loss moduli of the composites and the matrix, respectively. By substituting the measured values of E_c''/E_m'' into Eq. (11), one can obtain ϕ_e. The data showed that the effective volume fraction increased with increasing filler content and decreasing filler size.

In fact, most of the current theories valid for conventional composite systems assume an isotropic and homogeneous matrix filled with spherical particles. In this context, Young's moduli of the composites should be independent of the dimension of the dispersed phase. The local stresses in these composites under load are only dependent on the ratio of the distance between the particles of the dispersed phase and the dimensions of these particles. This ratio is often a constant in the case of a certain volume fraction of the dispersed phase and a certain spatial distribution of the particles. Evidently, this is not the case, especially when nanoparticles, characterized by high surface area, are incorporated. Therefore, the commonly accepted theories have to be modified in the case of nanocomposites, as shown by the above-mentioned works. Vollenberg and Heikens treated this problem in another way [172]. They clearly demonstrated the considerable dependence of Young's modulus on filler size (ranging from 35 nm to 100 μm); it was indicated that effects like adsorption of additives, particle size distribution, dewetting, and matrix morphology cannot explain the phenomenon well. The remaining explanation is that the solidification of the matrix is locally activated at a free filler particle surface, leading to a particular morphology of the polymer surrounding the filler particles, such as segmental orientation and improved packing in the neighborhood of the filler surface. For nanoparticles, the interparticle distance is so small that a homogeneous matrix material of a higher modulus polymer is reasonably assumed to be created, as supported by the results of annealing and solid-state nuclear magnetic resonance experiments [211].

It is also of importance to note that the formation of a relatively compliant layer at the interface tends to hinder a complete stress transfer under low stress and thus masks the stiffness of the nanoparticles. Table 11 gives evidence of such a masking effect provided by grafting polymers (PBA, PVA, and PEA, for example) on various nanoparticles. In this or a similar case, the above-discussed relationship between composite modulus and effective volume fraction of the interphase, consisting of an immobilized matrix polymer, is no longer applicable.

To avoid complicated geometric effects in composite materials, Shang and co-workers proposed a thermodynamic model based on the work of adhesion [170]. It was hypothesized that the interfacial bond strength can be quantified

by this parameter. The rationality of this assumption lies in the breakdown of the interface interaction required to do certain work. Accordingly, Young's modulus of a composite at a given filler loading is written as a function of the work of adhesion, W_a:

$$E_c/E_{co} = \exp\{-K_{E_c}[(1/W_a) - (1/W_{ao})]\} \qquad (12)$$

where E_{co} is the modulus of the composite with the weakest interfacial bond, W_{ao} is the work of adhesion of the composite with the weakest interfacial bond, and K_{E_c} is a constant, independent of the filler volume content and experimentally determined from the slope of log E_c versus $1/W_a$. When the Young's modulus of a composite with its corresponding W_a is known, Eq. (12) can be used to predict the Young's modulus of another composite with another W_a but with the same filler content. Results from SiO_2/EVA composites verified that Eq. (12) leads to good agreement for both, composites filled with micro- and nanosized particles. Furthermore, the data indicated that the composites that incorporated smaller particles have a higher K_{E_c} value. Since the K_{E_c} values show how the Young's modulus changes with W_a, the dependence of K_{E_c} on particle size proves that the effect of the interfacial bond on Young's modulus depends on filler particle size.

4.2.2. Strength and Fracture Toughness

The theories for the strength properties of filled systems are less developed than those for the modulus. In particular, only a few reports deal with the exploration of fracture mechanisms of nanocomposites where changes in the nature of crack propagation are believed to be one of the main mechanisms responsible for the strengthening effect [212]. In the case of poor bonding between the filler particles and the matrix, the tensile strength of a composite is usually reduced, with the filler content ϕ_f following a power law [213–216],

$$\sigma_c = \sigma_m(1 - \alpha\phi_f^n) \qquad (13)$$

where σ_c and σ_m are the strengths of the composite and the matrix, respectively; the constants α and n depend on the shape of the particles and their arrangement in the composite. This means that the strength of the composite cannot be greater than that of the unfilled version because the filler particles do not bear any fraction of the external load. However, this contradicts the results of the nanocomposites described in previous sections.

When the bonding between fillers and the matrix is strong enough, as suggested by Jancar and co-workers [217], the tensile strength of a composite can be 1.33 times as high as that of the matrix polymer:

$$\sigma_c/\sigma_m = 1 + 0.33F(c)\phi_f^2 \qquad (0 < F(c)\phi_f^2 \leq 1) \qquad (14)$$

$$\sigma_c/\sigma_m = 1.33 \qquad (1 < F(c)\phi_f^2) \qquad (15)$$

where $F(c)$ is proportional to the average yielded area per particle. In this case, an overlapping of stress concentration fields around the particulate fillers is responsible for shear yielding of the matrix [218]; the latter can be considered the dominant energy dissipation mechanism. Figure 32 illustrates good regression results in terms of Eqs. (14) and (15), demonstrating that a "perfect" adhesion is associated with the PP-based nanocomposites when $\phi_f > 0.65$ vol.% [3]. It should be noted that the equilibrium relative strength is about 1.15 but not 1.33, which can be attributed to the low ductility of the matrix polymer employed. When the filler content exceeds the critical value of 0.65 vol.% (i.e., $F(c)\phi_f^2 = 1$), the corresponding $F(c)$ can be estimated to be around 2.4×10^4. Considering the fact that $F(c)$ characterizes the rate of increase in connectivity of yielded microzones around individual particles [75], one might conclude that the reinforcing effect generated by the nano-SiO_2 particles here is far superior to that induced by conventional particles. For example, the $F(c)$ values for $CaCO_3$ (10 μm in diameter)-filled PP composites range only from 3 to 5 [217]. In addition, the greater value of $F(c)$ is also indicative of a strong interfacial interaction in the nanocomposites.

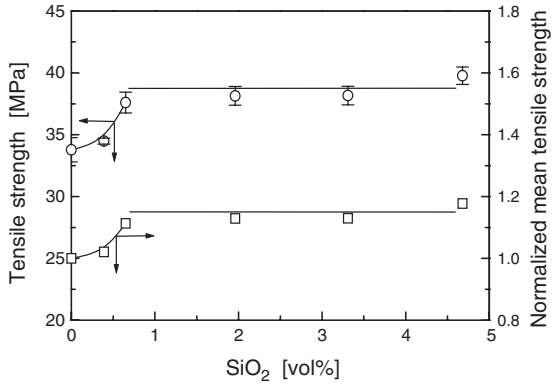


Figure 32. Tensile strength of SiO_2-g-PS/PP nanocomposites as a function of SiO_2 content. The lines represent the results of curve fitting in terms of Eqs. (14) and (15). Reprinted with permission from [3], M. Z. Rong et al., *Polymer* 42, 3301 (2001). © 2001, Elsevier Science.

Pukanszky et al. developed, alternatively, a semiempirical correlation for describing the composition dependence of the tensile yield stress in heterogeneous polymer systems [219]:

$$\sigma_c/\sigma_m = (1 - \phi_f)\exp(B\phi_f)/(1 + 2.5\phi_f) \qquad (16)$$

where B is a parameter related to the components' interaction. It is given by

$$B = (1 + A_f\rho_f l)\ln(\sigma_i/\sigma_m) \qquad (17)$$

where A_f and ρ_f are the specific surface area and the density of the filler, and l and σ_i are the thickness and yield stress of the interphase. By using Eq. (16) and the measured mechanical performance, Fekete et al. calculated a series of B values [220]. It was shown that the correlation of B versus A_f is linear up to a specific surface area of about 7–8 m^2/g as predicted by Eq. (17), but it becomes independent of A_f at higher values. The deviation from the prediction was attributed to the formation of particle agglomerates. With agglomeration, the surface area for interphase formation no longer changes with increasing A_f. Clearly, an examination of the linear relation between B and A_f based on Eq. (17) can reveal the homogeneity of the nanocomposites of interest.

Shang et al. also took the interfacial interaction into consideration when discussing the tensile strength of SiO_2/EVA composites [170]. Similar to the description of the modulus in Eq. (12), the dependence of tensile strength on the work of adhesion can be written as

$$\sigma_c/\sigma_{co} = \exp\{-K_{\sigma_c}[(1/W_a) - (1/W_{ao})]\} \quad (18)$$

where σ_c is the tensile strength of a composite with W_a at the interface. σ_{co} is the tensile strength of the composite with the weakest interfacial bond W_{ao}. When the tensile strength of a composite and its corresponding W_a are known, Eq. (18) can predict the tensile strength of another composite of a certain filler loading, even if the fillers have different W_a values. K_{σ_c} is the slope of log σ_c versus $1/W_a$ and depends on both silica particle size and volume fraction. The experimental results reflected that the K_{σ_c} values increased as the filler content increased, especially for the composites with coarser particles. It was thus suggested that for the tensile strengths of the nanosilica composites, the interfacial bond strength has a more pronounced effect on the composites filled with larger filler particles than on the composites with fine particles. The effect of W_a on the tensile strength is dependent on both particle size and volume fraction, but the Young's modulus depends only on filler size.

Sumita et al. observed that the shear yield stress for the composites filled with nanoparticles increases with increasing filler content and with decreasing filler size, whereas the value for the composites filled with micron-sized particles decreases with filler content [155, 156]. Considering that the dependence of reinforcing and antireinforcing effects on filler size and volume fraction can be explained by the dispersion strength theory, Sumita and co-workers proposed a modified equation as follows:

$$\tau_c = \tau_m(1 - \phi_f^{2/3}) + Gb/\{dk(d)[(4\pi/3\phi_f)^{1/3} - 2]/2\} \quad (19)$$

where τ_c and τ_m are the shear yield stresses of a composite and its matrix, G is the shear modulus, b is the Burger's vector, d is the diameter of a particle, and $k(d)$ is an aggregation parameter. It suggests that under the same filler content, smaller nanoparticles result in a more remarkable reinforcing effect. For the micron-sized particle-filled composites, the second item of Eq. (19) is negligibly small. The experimental results of nano-SiO_2 incorporated PP and nylon 6 exhibited a good agreement with the calculation with Eq. (19) [155, 156].

Xiong et al. concluded that the influence of particle size on the composite strength can be depicted more clearly by a simple empirical expression. The authors studied the mechanical performance of ultrafine Al_2O_3/PS composites prepared through bulk polymerization in the presence of filler particles [221]:

$$\sigma_c = 35\{1 + [2.71/(1 + R)]\} \quad (20)$$

where the units of σ_c and R are MPa and μm, respectively. The authors showed that particles with radii smaller than 0.5 μm can bring about significant reinforcing effects.

Based on the methodology of structural elements, Ovchinskii et al. developed a computer simulation model and algorithms for simulating the processes of failure of fibrous nanocomposites [212]. An account was made of the scale effect of the strength of the superthin fibers, which was associated with a change in not only their lengths but also their diameters. Analysis of the modeling results showed that the increase in the strength of the examined composites was obtained by employing superthin fibers, which was more efficient than increasing the fiber volume content.

The fracture toughness of $CaCO_3$ (70 nm)-filled PP was studied by Levita et al. [142]. It was found that the critical stress intensity factor remained almost unchanged in the lower filler concentration range (5 vol.%). Different behaviors were observed at higher loadings, depending on the surface treatment. There was a monotonic decrease in the case of untreated fillers, whereas for modified fillers a well-defined maximum at 10 vol.% was observed. As a coupling agent, stearic acid was more efficient than titanate. No difference between the two treatments was found at higher concentrations. Usually the fracture properties of particulate composites can be explained by the model of crack pinning [222]. It is assumed that the crack front is characterized by a line tension (energy per unit length), so that lengthening this line increases the energy associated with the crack. When the crack encounters an array of impenetrable particles it is arrested (pinning) and a higher tension is required to release the crack. Lange and Redford related the fracture energy of a composite, G_c, with the line tension, T [223]:

$$G_c = G_m + 3\phi_f T/2d(1 - \phi_f) \quad (21)$$

where G_m is the fracture energy of the matrix. To account for the influence of particle size, an empirical factor, $F(d)$ $(0 < F(d) < 1)$, has been introduced [224]. At low concentration Eq. (21) thus becomes

$$G_c = G_m + 3\phi_f F(d)T/2d \quad (22)$$

By examining the $F(d)T$ values as a function of particle size, Levita et al. pointed out that $F(d)T$ is very small for finely sized particles. That is, the pinning contribution proved to be negligible. A particle can only interact with a crack if its size is greater than a certain length that characterizes the crack. Instead, interfacial debonding that reduces the local stress was attributed to the main factor responsible for the toughening effect.

Studies on the filler loading dependence of the toughness of SiO_2/PP nanocomposites showed that the maxima in impact strength and fracture toughness appeared already at very low filler loadings (<1 – 3 vol.%) [107]. This is the opposite of the optimum loading conditions known for conventional particle-filled systems. Although the typical size of the modified nanoparticles at the above-mentioned filler content range varies between 100 and 150 nm (as found by TEM observation), the distances between individual particles and therefore the interparticle ligament thicknesses are still much larger than those found in toughness-optimized conventional particle-filled systems [225]. Therefore, the considerable toughening effect perceived at such a low filler loading in the nanocomposites implies that the mechanisms involved should be different from those used as a basis for the single-percolation concept of Wu [226], which has been proved to be well applicable to conventional particulate-filled polymer composites.

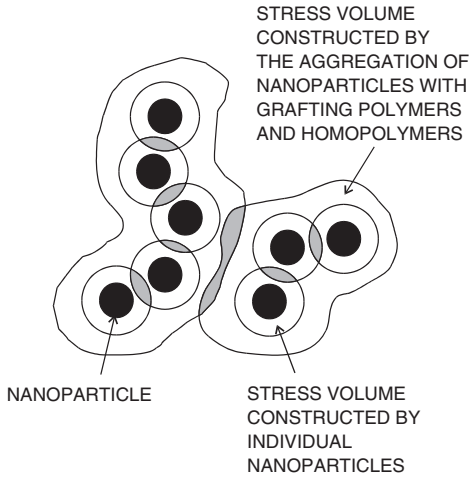


Figure 33. Hypothetical model of a nanocomposite with superposed stress volumes.

To overcome this problem, a hypothetical model of double percolation was suggested [3]. It was assumed that a double percolation, characterized by the appearance of connected shear-yielded networks throughout the composite, might be responsible for the performance enhancement at low nanofiller loading. This means one is dealing with (i) a percolation of shear yielded zones inside the dispersed phases (i.e., SiO_2-g-PS agglomerates, consisting of nano-SiO_2, the grafting PS on SiO_2, and the homopolymer derived from the grafting styrene monomer) due to the superposition of stress volumes around the nanoparticles, and (ii) a percolation of shear-yielded zones throughout the matrix resin due to the superposition of stress volumes around the dispersed agglomerates (Fig. 33). Provided these two percolation processes take place almost simultaneously, the eventual percolation threshold should be equal to the product of the individual percolation thresholds, in accordance with the scaling theory [227]. As a result, a rather low filler content is sufficient to bring about a significant improvement of the nanocomposites' performance. In fact, the concept of a so-called double percolation or multiple percolation has been successfully used in designing very low filler-loaded conductive polymer composites [228]. It should also be applicable to the description of critical transitions in other disordered systems with randomly geometric structure. In addition to the quantitative analysis provided in [3], fractography of tensile-fractured surfaces of the nanocomposites also give supporting evidence, as illustrated by the closely packed concentric matrix-fibrillated circles around particle-like objects (Fig. 34).

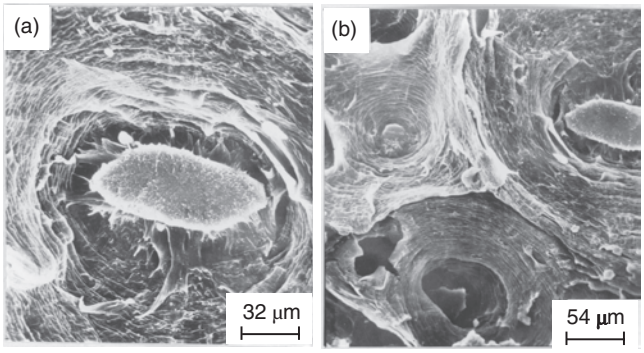


Figure 34. Scanning electron microscopy (SEM) micrographs of a tensile-fractured surface of SiO_2-g-PS/PP nanocomposites (content of SiO_2= 1.96 vol.%). Reprinted with permission from [107], M. Z. Rong et al., *Polymer* 42, 167 (2001). © 2001, Elsevier Science.

5. CONCLUSION

Nanocomposites represent a new prospective branch in the huge field of polymer materials science and technology. Though they have not yet been thoroughly studied, they show promising results in many aspects. To improve their competitiveness in the market, direct mixing of nanoparticles and polymers by existing blending techniques proved to be a possible breakthrough as long as the fillers are suitably pretreated. In contrast to the composites filled with microparticles, nanoparticles might permit simultaneous enhancement of the modulus, strength, toughness, and thermal deformation temperature of polymers without affecting the thermoformability under a rather low filler loading [3]. In addition, the rheological performance of the melts of the nanosystem is similar to those of the unfilled versions. As a result, the density, cost, and gloss of the final products are nearly not influenced, providing broad applicability in a variety of processing and manufacturing.

To truly ascertain the present and future potential applications of polymer nanocomposites, considerably more research and development needs to be done. The following issues in particular have to be borne in mind:

- What are the theoretical upper limits of the mechanical properties of a polymer nanocomposite?

So far, the reported increment in strength and toughness of nanocomposites (including intercalated and exfoliated nanocomposites) relative to the properties of the matrices is less than one order of magnitude, regardless of how the composites were prepared. It strongly indicates the necessity of establishing a framework that is able to quantitatively predict the reinforcing effects of nanoparticles as well as of new compounding techniques. These are the prerequisites for further development of the nanocomposites.

- How does the interphase influence the mechanical performance of a polymer nanocomposite?

In fact, this question is closely related to the previous one. When the reinforcing fillers reach a nanoscopic size, the mechanical properties of the nanocomposites should be governed almost entirely by the interface, and the bulk properties play a secondary role [5]. As the nanofillers might have the same size as the segments of the surrounding polymer chains, the classic models are no longer valid.

- What kind of surface treatment is most effective for purposes of separating nanoparticle agglomerates and improving filler/matrix interfacial adhesion?

It seems to be somewhat difficult to simultaneously meet both demands. Actually, no matter what methods have been used to date, a uniform dispersion of nanoparticles on a

nanometer scale is hardly obtained throughout the composites when it is done by the direct addition of premade nanoscale fillers to a thermoplastic matrix. Therefore, in addition to a continuous study of novel dispersion techniques and compounding facilities, research interests should also focus on developing surface treatments capable of giving full play to the reinforcing effect of nanoparticulate fillers, even though a homogeneous dispersion in the matrix is not available.

- What is the optimum phase morphology of nanocomposites, and how can this morphology be obtained by choosing proper processing conditions (including surface treatments of nanoparticles and mixing variables)?

Self-organization and self-assembly have been successfully adopted in the construction of well-defined discrete nanosystems, with a physical oriented performance toward certain functions [19, 229, 230]. Comparatively speaking, the effect of the spatial distribution of nanoparticles or nanoparticle agglomerates on the mechanical properties of nanocomposites has not yet been understood, to say nothing of the effects of the adjustment of the microstructural arrangement.

- Are there any relationships between the species of the nanoparticles and the improvement in mechanical performance of the nanocomposites?

The study of PP-based nanocomposites demonstrated that SiO_2 and $CaCO_3$ have similar strengthening and toughening effects [153]. It seems that a proper interfacial bonding, which facilitates (i) stress transfer between the matrix and the nanoparticles and (ii) plastic deformation of the matrix, is a very important issue. If this is the case, however, it is somewhat difficult to understand the positive synergistic effect provided by ultrafine $CaCO_3$ and talc [231], which cannot perform well when used individually in reinforcing HDPE.

- What are the applications for which only nanocomposites are competent?

The superiority of a new material should be expressed by its unique technical importance. For example, Al_2O_3 is not a suitable filler in the microscale particulate form for wear-resisting composites because of its angularity and thus its propensity to damage the counterface. However, the material in the nanoscale particulate form has much lower angularity and, therefore, is not that abrasive [154].

To have an idea of the prospects of the nanocomposites, a brief survey of the history of the development of intrinsically conducting polymers might be enlightening. Much of the early research on conducting polymers was spurred by their potential application as replacements for existing metals and semiconductors. However, with many years of intensive development activity, it seems that many of the proposed applications for these materials were rather optimistic and have not yet become possible. With growing experience, a number of more realistic applications have now emerged that tend to exploit the novel features associated with conducting polymers rather than those properties that are readily obtainable in traditional materials [232]. Similarly, the future commercial success of polymer nanocomposites might also require a careful consideration of other technological and economic aspects besides their mechanical role.

GLOSSARY

Agglomerates Weakly bonded particle clusters.

Aggregates Strongly bonded particle clusters.

Compatibility Miscibility on a molecular scale.

Composites A material consisting of two or more components, which largely retain their identities and yet result in many properties that cannot be achieved with either of the constituents acting alone.

Copolymer A polymer in which chemical combination exists in the main chain between two polymers $[A]_n$ and $[B]_n$. A copolymer can be a block copolymer [AAA...] [BB...] or a random copolymer ABAABAB, the latter having no long sequences of A or B units.

Crystallization of polymers The macromolecular chains fold in a regular manner forming ordered structure.

Dispersion Break up of agglomerated particles and separation of the resulting fragments to a point where reagglomeration will not occur.

Failure In a tensile stress-strain experiment, the sample is elongated until it breaks. Stress-strain studies are usually relatively slow, of the order of mm/s. Impact strength measures the material's resistance to a sharp blow, typically m/s. In both stress-strain and impact studies, energy is absorbed within the sample by viscoelastic deformation of the polymer chains, and finally by the creation of new surface areas.

Glass transition A transition describing the change between glassy and rubbery states in polymers. At the glass transition temperature, the amorphous portions of a polymer soften. Carried out under ideal conditions, glass transition is a type of second-order transition.

Graft Long macromolecular chains of a polymer are chemically attached to the base material.

Initiation Radical chain polymerization is a chain reaction consisting of a sequence of three steps-initiation, propagation and termination. The initiation step is considered the production of free radicals by any one of a number of reactions. The initiation involves the addition of this radical to the first monomer molecule to produce the chain initiating species.

Interface A surface formed by a common boundary of filler and matrix that is in contact with and maintains the bond in between. A polymer composite interface consists of a polymer and a nonpolymer. Usually the nonpolymer is a solid phase such as glass, carbon, or boron fibers, steel, or particulates such as calcium carbonate or titanium dioxide. In such a case the polymer cannot interdiffuse into the nonpolymer surface but may bond and adhere to it.

Monomers Small molecules that combine with each other to form polymer molecules.

Polymer A long molecule containing atoms held together by primary covalent bonds along the molecule. When one speaks of polymers, one is concerned with materials whose molecular weights may reach into the millions.

Polymerization Reactions by which the monomers combine building up macromolecules.

Thermoplastics Linear or branched polymers, that is, uncrosslinked ones, which soften, melt and flow upon heating. Almost all of the polymers produced by chain polymerization are thermoplastics. There are only a relatively few such polymers which are crosslinked.

Wetting Polymer wets the filler and squeezes into its void spaces, such that losse filler particles disappear and air introduced into the compound by entrapment in the filler agglomerates is replaced.

Young's modulus A parameter characterizing stiffness of materials, which relates the stress to the strain for uniaxial tensile or compressive loading.

ACKNOWLEDGMENTS

The authors are grateful for the support of the Volkswagen-Stiftung (Grant I/76645) and the Deutsche Forschungsgemeinschaft (DFG FR675/40-1) for the cooperation between the German and the Chinese institutes on the topic of nanocomposites. Further thanks are due to the National Natural Science Foundation of China (Grant 50133020), the Key Program of the Ministry of Education of China (Grant 99198), and the Team Project of the Natural Science Foundation of Guangdong, China.

REFERENCES

1. R. Dagani, *Chem. Eng. News* 70, 18 (1992).
2. R. Dagani, *Chem. Eng. News* 77, 25 (1999).
3. M. Z. Rong, M. Q. Zhang, Y. X. Zheng, H. M. Zeng, and K. Friedrich, *Polymer* 42, 3301 (2001).
4. A. D. Pomogailo, *Russ. Chem. Rev.* 69, 53 (2000).
5. E. Reynaud, C. Gauthier, and J. Perez, *Rev. Metallurgie* 96, 169 (1999).
6. G. Carotenuto, L. Nicolais, X. Kuang, and Z. Zhu, *Appl. Compos. Mater.* 2, 385 (1995).
7. R. H. Norman, "Conductive Rubbers and Plastics." Elsevier, Amsterdam, 1970.
8. H. Zeng, Z. Zhang, M. Zhang, J. Xu, N. Jian, and K. Mai, *J. Appl. Polym. Sci.* 54, 541 (1994).
9. A. B. R. Mayer and J. E. Mark, *Polym. Bull.* 37, 683 (1996).
10. A. B. R. Mayer and J. E. Mark, *Colloid Polym. Sci.* 275, 333 (1997).
11. S. W. Huang, K. G. Neoh, E. T. Kang, H. S. Han, and K. L. Tan, *J. Mater. Chem.* 8, 1743 (1998).
12. K. G. Neoh, K. K. Tan, P. L. Goh, S. W. Huang, E. T. Kang, and K. L. Tan, *Polymer* 40, 887 (1999).
13. L. Zhang, C. Xie, and X. Zhu, *J. Alloys Compd.* 211/212, 390 (1994).
14. M. Z. Rong, M. Q. Zhang, Y. X. Zheng, and H. M. Zeng, Chinese Patent Application CN99116017, 1999.
15. D. Y. Godovski, *Adv. Polym. Sci.* 119, 79 (1995).
16. M. D. Butterworth, S. A. Bell, S. P. Armes, and A. W. Simpson, *J. Colloid Interface Sci.* 183, 91 (1996).
17. M. Wan and J. Li, *J. Polym. Sci., Polym. Chem.* 36, 2799 (1998).
18. B. Z. Tang, Y. Gang, J. W. Y. Lam, B. Li, X. Jing, X. Wang, F. Wang, A. B. Pakhomov, and X. X. Zhang, *Chem. Mater.* 11, 1581 (1999).
19. M. Z. Rong, M. Q. Zhang, H. Liu, and H. M. Zeng, *Polymer* 40, 6169 (1999).
20. W. A. Wampler, K. Rajeshwar, R. G. Pethe, R. C. Hyer, and S. C. Sharma, *J. Mater. Res.* 10, 1811 (1995).
21. R. Gangopadhyay and A. De, *Chem. Mater.* 12, 608 (2000).
22. M. Yoshida, M. Lal, N. D. Kumar, and P. N. Prasad, *J. Mater. Sci.* 32, 4047 (1997).
23. C. J. Wung, Y. Pang, P. N. Prasad, and F. E. Karasz, *Polymer* 32, 605 (1991).
24. N. M. Lawandy, R. M. Balachandran, A. S. L. Gomes, and E. Sauvain, *Nature* 368, 436 (1994).
25. H. Yoneyama, *Adv. Mater.* 5, 394 (1993).
26. L. L. Beecroft and C. K. Ober, *Chem. Mater.* 9, 1302 (1999).
27. X. Chen and K. E. Gonsalves, *J. Mater. Res.* 12, 1274 (1997).
28. G. Cao, M. E. Garcia, M. Aleala, L. F. Burgess, and T. E. Mallouk, *J. Am. Chem. Soc.* 114, 7574 (1992).
29. Y. Chen, H. L. W. Chan, N. M. Hui, Y. W. Wong, and C. L. Choy, *Sens. Actuators, A* 69, 156 (1998).
30. J. Burdon and P. Calvert, *Mater. Res. Soc. Symp. Proc.* 286, 315 (1993).
31. Q. Wang, J. Xu, W. Shen, and Q. Xue, *Wear* 209, 316 (1997).
32. J. Luo, J. J. Lannutti, and R. R. Seghi, *Dent. Mater.* 14, 29 (1998).
33. P. Calvert, *Mater. Sci. Eng.* C1, 69 (1994).
34. S. Zhang and K. E. Gonsalves, *Mater. Sci. Eng.* C3, 117 (1995).
35. R. Zeng, Ph.D. Thesis, Zhongshan University, 2000.
36. Y. Kojima, A. Usuki, M. Kawasumi, A. Okada, Y. Fukushima, T. Kurauchi, and O. Kamigaito, *J. Polym. Sci., Polym. Chem.* 31, 1755 (1993).
37. E. P. Giannelis, *Adv. Mater.* 8, 29 (1996).
38. H.-L. Tsai, J. L. Schindler, C. R. Kannewurf, and M. G. Kanatzidis, *Chem. Mater.* 9, 875 (1997).
39. A. D. Pomogailo, *Plat. Met. Rev.* 38, 60 (1994).
40. M. Sumita, T. Okuma, K. Miyasaka, and K. Ishikawa, *J. Appl. Polym. Sci.* 27, 3059 (1982).
41. S. W. Shang, J. W. Williams, and K.-J. M. Soderholm, *J. Mater. Sci.* 27, 4949 (1992).
42. Y. Dirix, C. Bastiaansen, W. Caseri, and P. Smith, *J. Mater. Sci.* 34, 3859 (1999).
43. Y. Nakao, *J. Colloid Interface Sci.* 171, 386 (1995).
44. C. H. Griffiths, M. P. Ohoro, and T. W. Smith, *J. Appl. Phys.* 50, 7108 (1979).
45. K. E. Gonsalves, G. Carlson, X. Chen, S. K. Gayen, R. Perez, and M. Jose-Yacaman, *Polym. Mater. Sci. Eng.* 73, 298 (1995).
46. H. Liu, M.Sc. Thesis, Zhongshan University, 1997.
47. C. Zilg, R. Thomann, R. Muelhaupt, and J. Finter, *Adv. Mater.* 11, 49 (1999).
48. E. P. Giannelis, *JOM* 44, 28 (1992).
49. B. M. Novak, *Adv. Mater.* 5, 422 (1993).
50. X. Xu, Y. Li, X. Ge, H. Wu, and Z. Zhang, *Mater. Lett.* 37, 354 (1998).
51. M. Hussain, A. Nakahira, S. Nishijima, and K. Niihara, *Mater. Lett.* 26, 299 (1996).
52. P. M. Ajayan, L. S. Schadler, C. Giannaris, and A. Rubio, *Adv. Mater.* 12, 750 (2000).
53. H. Ge and J. Wang, *Fiber Reinforced Plast. Compos.* 3, 13 (1999) (in Chinese).
54. T. Adebahr, C. Roscher, and J. Adam, *Eur. Coatings J.* 4, 144 (2001).
55. B. Pukanszky and E. Fekete, *Adv. Polym. Sci.* 139, 109 (1999).
56. W. B. Russel, D. A. Saville, and W. R. Schowalter, "Colloidal Dispersions." Cambridge University Press, Cambridge, UK, 1989.
57. R. N. Rothon, *Adv. Polym. Sci.* 139, 67 (1999).
58. B. Pukanszky and E. Fekete, *Polym. Polym. Compos.* 6, 313 (1998).
59. K. Kendall, in "Tribology in Particulate Technology" (B. J. Briscoe and M. J. Adams, Eds.), p. 154. Adam Hilger, Bristol, 1987.
60. M. J. Adams and B. Edmondson, in "Tribology in Particulate Technology" (B. J. Briscoe and M. J. Adams, Eds.), p. 135. Adam Hilger, Bristol, 1987.
61. B. Pukanszky, in "Polypropylene: An A-Z Reference" (J. Karger-Kocsis, Ed.), p. 574. Kluwer, Dordrecht, the Netherlands, 1999.

62. E. P. Plueddemann, "Silane Coupling Agents." Plenum Press, New York, 1982.
63. E. Fekete, B. Pukanszky, A. Toth, and I. Bertoti, *J. Colloid Interface Sci.* 135, 200 (1990).
64. E. Papirer, J. Schultz, and C. Turchi, *Eur. Polym. J.* 20, 1155 (1984).
65. T. Nakatsuka, H. Kawasaki, and K. Itadani, *J. Appl. Polym. Sci.* 27, 259 (1982).
66. H. L. Pan, *Ind. Inorg. Salt* 1, 24 (1996) (in Chinese).
67. J. G. Zhou, B. S. Li, X. X. Huang, and J. K. Guo, *J. Inorg. Mater.* 11, 237 (1996) (in Chinese).
68. T. Yoshihara, *Int. J. Adhes. Adhes.* 19, 353 (1999).
69. H. Li, *Plastics* 28, 25 (1999) (in Chinese).
70. K. Yoshinaga, T. Yokoyama, Y. Sugawa, H. Krakawa, N. Enomoto, H. Nishida, and M. Komatsu, *Polym. Bull.* 28, 663 (1992).
71. M. Hasegawa, K. Arai, and S. Saito, *J. Polym. Sci., Polym. Chem.* 25, 3117 (1987).
72. H. Du, Z. Q. Liu, F. Q. Liu, T. J. Li, X. Y. Tang, W. Q. Xu, and Y. Wu, *Chem. J. Chin. U.* 18, 1565 (1997) (in Chinese).
73. H. B. Liu, *Chemistry* 10, 44 (1997) (in Chinese).
74. K. E. Gonsalves and X. Chen, *J. Mater. Res.* 12, 1274 (1997).
75. J. Jancar, in "Handbook of Polypropylene and Polypropylene Composites" (H. G. Karian, Ed.), p. 367. Dekker, New York 1999.
76. J. P. Blitz and C. B. Little, "Fundamental and Applied Aspects of Chemically Modified Surfaces." Royal Society of Chemistry, Cambridge, UK, 1999.
77. E. F. Vansant, P. van der Voort, and K. C. Vrancken, "Characterization and Chemical Modification of the Silica Surface." Elsevier, Amsterdam, 1995.
78. V. M. Gunko, E. F. Voronin, E. M. Pakhlov, V. I. Zarko, V. V. Turov, N. V. Guzenko, R. Leboda, and E. Chibowski, *Colloids Surf., A* 166, 187 (2000).
79. M. I. Baraton, F. Chancel, and L. Merhari, *Nanostruct. Mater.* 9, 319 (1997).
80. Ph. Espiard and A. Guyot, *Polymer* 36, 4391 (1995).
81. M. L. C. M. Oosterling, A. Sein, and A. J. Schouten, *Polymer* 33, 4394 (1992).
82. E. Schomaker, A. J. Zwarteveen, G. Challa, and M. Capka, *Polym. Commun.* 29, 158 (1988).
83. N. Tsubokawa, A. Kogure, and Y. Sone, *J. Polym. Sci., Polym. Chem.* 28, 1923 (1990).
84. N. Tsubokawa, H. Ishida, and K. Hashimoto, *Polym. Bull.* 31, 456 (1993).
85. N. Tsubokawa and H. Ishida, *J. Polym. Sci., Polym. Chem.* 30, 2241 (1992).
86. S. Hayashi, Y. Takeuchi, M. Eguchi, T. Iida, and N. Tsubokawa, *J. Appl. Polym. Sci.* 71, 1491 (1999).
87. N. Tsubokawa, A. Kogure, K. Maruyama, Y. Sone, and M. Shimomura, *Polym. J.* 22, 827 (1990).
88. N. Tsubokawa, Y. Shirai, H. Tsuchida, and S. Handa, *J. Polym. Sci., Polym. Chem.* 32, 2327 (1994).
89. G. Boven, M. L. C. M. Oosterling, G. Challa, and A. J. Schouten, *Polymer* 31, 2377 (1990).
90. N. Tsubokawa, K. Maruyama, Y. Sone, and M. Shimomura, *Polym. J.* 21, 475 (1989).
91. Y. Shirai, K. Kawatsura, and N. Tsubokawa, *Prog. Org. Coat.* 36, 217 (1999).
92. N. Tsubokawa, Y. Shirai, and K. Hashimoto, *Colloid Polym. Sci.* 273, 1049 (1995).
93. O. W. Webster, *Science* 251, 887 (1991).
94. T. von Werne and T. E. Patten, *J. Am. Chem. Soc.* 123, 7497 (2001).
95. N. Tsubokawa and S. Yoshikawa, *J. Polym. Sci., Polym. Chem.* 33, 551 (1995).
96. Q. Liu, J. R. D. Wijn, K. D. Groot, and C. A. V. Blitterswijk, *Biomaterials* 19, 1067 (1998).
97. M. L. Green, W. E. Rhine, P. Calvert, and H. K. Bowen, *J. Mater. Sci. Lett.* 12, 1425 (1993).
98. K. Fukano and E. Kageyama, *J. Polym. Sci., Polym. Chem.* 13, 1309 (1975).
99. K. Fukano and E. Kageyama, *J. Polym. Sci., Polym. Chem.* 13, 1325 (1975).
100. K. Fukano and E. Kageyama, *J. Polym. Sci., Polym. Chem.* 14, 1743 (1976).
101. K. Fukano and E. Kageyama, *J. Polym. Sci., Polym. Chem.* 15, 65 (1977).
102. K. Fukano and E. Kageyama, *J. Polym. Sci., Polym. Chem.* 14, 2193 (1976).
103. K. Fukano and E. Kageyama, *J. Polym. Sci., Polym. Chem.* 14, 2183 (1976).
104. K. Fukano and E. Kageyama, *J. Polym. Sci., Polym. Chem.* 14, 23 (1976).
105. K. Fukano and E. Kageyama, *J. Polym. Sci., Polym. Chem.* 14, 1031 (1976).
106. B. Jiang and G. L. Huang, *J. Radiat. Res. Radiat. Process.* 14, 92 (1996) (in Chinese).
107. M. Z. Rong, M. Q. Zhang, Y. X. Zheng, H. M. Zeng, R. Walter, and K. Friedrich, *Polymer* 42, 167 (2001).
108. R. Herzig and W. E. Baker, *J. Mater. Sci.* 28, 6531 (1993).
109. G. D. Parfitt, "Dispersion of Powders in Liquids." Applied Science, London, 1973.
110. D. S. Parker, W. J. Kaufman, and D. Jenkins, *J. San. Eng. Div. Proc. Am. Soc. Civ. Eng.* 98(SA1), 79 (1972).
111. S. P. Rwei, I. Manas-Zloczower, and D. L. Feke, *Polym. Eng. Sci.* 30, 701 (1990).
112. S. P. Rwei, I. Manas-Zloczower, and D. L. Feke, *Polym. Eng. Sci.* 31, 558 (1991).
113. S. P. Rwei, I. Manas-Zloczower, and D. L. Feke, *Polym. Eng. Sci.* 32, 130 (1992).
114. J. D. Pandya and L. A. Spielman, *J. Colloid Interface Sci.* 90, 517 (1982).
115. S. V. Kao and S. G. Mason, *Nature* 253, 619 (1975).
116. R. L. Powell and S. G. Mason, *AIChE J.* 28, 286 (1982).
117. S. Shiga and M. Furuta, *Rubber Chem. Technol.* 58, 1 (1985).
118. Y. J. Lee, I. Manas-Zloczower, and D. L. Feke, *Polym. Eng. Sci.* 35, 1037 (1995).
119. H. Schubert, W. Herrmann, and H. Raumpf, *Powder Technol.* 11, 121 (1975).
120. H. Yamada, I. Manas-Zloczower, and D. L. Feke, *Powder Technol.* 92, 163 (1997).
121. H. Yamada, I. Manas-Zloczower, and D. L. Feke, *Rubber Chem. Technol.* 71, 1 (1998).
122. H. Yamada, I. Manas-Zloczower, and D. L. Feke, *Chem. Eng. Sci.* 53, 159 (1998).
123. F. Bohin, I. Manas-Zloczower, and D. L. Feke, *Rubber Chem. Technol.* 67, 602 (1994).
124. P. Levresse, I. Manas-Zloczower, D. L. Feke, Y. Bomal, and D. Bortzmeyer, *Powder Technol.* 106, 62 (1999).
125. J. Winkler, E. Klinke, and L. Dulog, *J. Coat. Technol.* 59, 35 (1987).
126. J. Winkler, E. Klinke, M. N. Sathyanarayana, and L. Dulog, *J. Coat. Technol.* 59, 45 (1987).
127. J. Winkler and L. Dulog, *J. Coat. Technol.* 59, 55 (1987)
128. R. Rothon, "Particulate-Filled Polymer Composites." Longman, Harlow, 1995.
129. J. E. Curry, in "Polymer Blends and Alloys" (M. J. Folkes and P. S. Hope, Eds.), p. 7. Blackie Academic & Professional, London, 1993.
130. H. Mack, *Plast. Eng.* 53, 53 (1997).
131. C. H. Suh and J. L. White, *J. Non-Newtonian Fluid Mech.* 62, 175 (1996).
132. J. W. Ess and P. R. Hornsby, *Plast. Rubber Process. Appl.* 8, 147 (1997).
133. M. Bories, M. A. Huneault, and P. G. Lafleur, *Int. Polym. Process.* 14, 234 (1999).

134. R. Huang, W. Xu, X. Zheng, L. She, B. Cai, and W. Fan, *Chin. Plast. Ind.* 25, 106 (1997) (in Chinese).
135. Y. Su, R. Huang, B. Cai, and W. Fan, *Chin. Plast.* 12, 22 (1998) (in Chinese).
136. W. Xu, R. Huang, B. Cai, and W. Fan, *Chin. Plast.* 13, 25 (1999) (in Chinese).
137. W. Xu, R. Huang, B. Cai, and W. Fan, *Chin. Plast.* 12, 30 (1998) (in Chinese).
138. W. Guo, D. Li, C. Su, S. Tang, and Z. Xu, *Chin. Plast. Ind.* 26, 10 (1998) (in Chinese).
139. K. Si and J. Tian, *Chin. Plast.* 4, 35 (1990) (in Chinese).
140. X. Ren, L. Bai, and G. Wang, *Chem. World* 41, 83 (2000) (in Chinese).
141. Z. Luo, R. Huang, A. Lu, B. Cai, and W. Fan, *Chin. Plast.* 13, 47 (1999) (in Chinese).
142. G. Levita, A. Marchetti, and A. Lazzeri, *Polym. Compos.* 10, 39 (1989).
143. P. W. M. Peters and H. Albertsen, *J. Mater. Sci.* 28, 1059 (1993).
144. G. Krekel, U. J. Zielke, K. J. Hunttinger, and W. P. Hoffman, *J. Mater. Sci.* 29, 3084 (1994).
145. H. E. H Meijer, R. W. Venderbosch, J. G. P. Goossens, and P. J. Lemstra, *High Perform. Polym.* 8, 133 (1996).
146. A. Saalbrink, A. Lorteije, and T. Peijs, *Compos. Part A, Appl. S.* 29A, 1243 (1998).
147. R. W. Venderbosch, T. Peijs, H. E. H. Meijer, and P. J. Lemstra, *Compos. Part A, Appl. S.* 27, 895 (1996).
148. S. C. Jana and S. Jain, *Polymer* 42, 6897 (2001).
149. X. Ren, L. Bai, G. Wang, and B. Zhang, *Chin. Plast.* 14, 22 (2000) (in Chinese).
150. J. Zhang, X. Wang, L. Lu, and X. Yang, *Chin. Plast.* 15, 24 (2001) (in Chinese).
151. X. Wang, R. Huang, C. Jin, H. Chen, and Y. Pu, *Chin. Plast.* 14, 34 (2000) (in Chinese).
152. S. Wu and P. Ye, *Eng. Plast. Appl.* 25, 12 (1997) (in Chinese).
153. M. Z. Rong, M. Q. Zhang, Y. X. Zheng, H. M. Zeng, R. Walter, and K. Friedrich, *J. Mater. Sci. Lett.* 19, 1159 (2000).
154. C. J. Schwartz and S. Bahadur, *Wear* 237, 261 (2000).
155. M. Sumita, Y. Tsukumo, K. Miyasaka, and K. Ishikawa, *J. Mater. Sci.* 18, 1758 (1983).
156. M. Sumita, T. Shizuma, K. Miyasaka, and K. Ishikawa, *J. Macromol. Sci., Phys.* B22, 601 (1983).
157. M. Sumita, T. Shizuma, K. Miyasaka, and K. Ishikawa, *Polym. Compos.* 7, 36 (1986).
158. M. Sumita, T. Ookuma, K. Miyasaka, and K. Ishikawa, *Colloid Polym. Sci.* 262, 103 (1984).
159. M. Sumita, H. Tsukihi, K. Miyasaka, and K. Ishikawa, *J. Appl. Polym. Sci.* 29, 1523 (1984).
160. C. J. T. Landry, B. K. Coltrain, M. R. Landry, J. J. Fitzgerald, and V. K. Long, *Macromolecules* 26, 3702 (1993).
161. E. V. Barrera, *JOM* 52, 38 (2000).
162. C. Cooper, D. Feigelshtein, and H. D. Wagner, in "Proceedings of the 13th International Conference on Composite Materials (ICCM-13)" (Y. Zhang, Ed.), Beijing, 2001, ID 1536.
163. C. Cooper, D. Ravich, D. Lips, J. Mayer, and H. D. Wagner, *Compos. Sci. Technol.* 62, 1105 (2002).
164. L. S. Schadler, S. C. Giannaris, and P. M. Ajayan, *Appl. Phys. Lett.* 73, 3842 (1998).
165. S. Hu, *Chin. Plast.* 13, 25 (1999) (in Chinese).
166. G. Carotenuto, L. Nicolais, X. Kuang, and Z. Zhu, *Appl. Compos.* 2, 385 (1995).
167. W. P. Hsu, Y. Rongchi, and E. Matijevic, *Dyes Pigm.* 19, 179 (1992).
168. G. Carotenuto, L. Nicolais, and X. Kuang, *Appl. Compos.* 3, 103 (1996).
169. D. Hamdoun, D. Ausserre, S. Joly, Y. Gallot, V. Cabuil, and C. Clinard, *J. Phys. II France* 6, 493 (1996).
170. S. W. Shang, J. W. Williams, and K.-J. M. Soederholm, *J. Mater. Sci.* 29, 2406 (1994).
171. S. W. Shang, J. W. Williams, and K.-J. M. Soederholm, *J. Mater. Sci.* 30, 4323 (1995).
172. P. H. T. Vollenberg and D. Heikens, *Polymer* 30, 1656 (1989).
173. M. S. P. Shaffer, X. Fan, and A. H. Windle, *Adv. Mater.* 11, 937 (1999).
174. M. S. P. Shaffer, X. Fan, and A. H. Windle, *Carbon* 36, 1603 (1998).
175. L. S. Schadler, K. O. Laul, R. W. Smith, and E. Petrovicova, *J. Therm. Spray Technol. 6*, 475 (1997)
176. E. Petrovicova, R. Knight, L. S. Schadler, and T. E. Twardowski, *J. Appl. Polym. Sci.* 77, 1684 (2000).
177. E. Petrovicova, R. Knight, L. S. Schadler, and T. E. Twardowski, *J. Appl. Polym. Sci.* 78, 2272 (2000).
178. T. Ishida and S. Tamaru, *J. Mater. Sci. Lett.* 12, 1851 (1993).
179. S. L. Namboodri, H. Zhou, A. Aning, and R. G. Kander, *Polymer* 35, 4088 (1994).
180. P. Hu, S. Fan, and J. Wan, *Eng. Plast. Appl.* 26, 1 (1998) (in Chinese).
181. C. Chen, S. Qian, F. Gong, Z. Wang, J. Jiang, and X. Yang, *Chin. J. Mater. Res.* 14, 334 (2000) (in Chinese).
182. J. Jurga, M. Nowicki, K. Bula, B. Susla, and S. S. Rejeibi, *Mol. Cryst. Liq. Cryst.* 354, 43 (2000).
183. H. Eilers, *Kolloid-Z*. 97, 313 (1941).
184. U. Zorll, *J. Appl. Polym. Sci.* 10, 1315 (1966).
185. Y. Nakao, *J. Chem. Soc., Chem. Commun.* 10, 826 (1993).
186. E. Kerner, *Proc. R. Soc. London, Ser. B* 69, 808 (1956).
187. G. C. Papanicolau and P. S. Theocaris, *Colloid Polym. Sci.* 257, 239 (1979).
188. K. Kendall and F. R. Sherliker, *Br. Polym. J.* 12, 85 (1980).
189. M. Kryszewski and G. W. Bak, *Acta Phys. Pol., A* 92, 1163 (1997).
190. J. K. Kim and Y. W. Mai, "Engineered Interfaces in Fiber Reinforced Composites." Elsevier, Amsterdam, 1998.
191. K. Kendall and F. R. Sherliker, *Brit. Polym. J.* 12, 111 (1980).
192. M. Q. Zhang, M. Z. Rong, H. M. Zeng, S. Schmitt, B. Wetzel, and K. Friedrich, *J. Appl. Polym. Sci.* 80, 2218 (2001).
193. V. F. Shumsky, E. G. Privalko, V. M. Karaman, V. P. Privalko, R. Walter, K. Friedrich, M. Q. Zhang, and M. Z. Rong, *Adv. Comps. Lett.* 10, 191 (2001).
194. M. S. Kunz, K. R. Shull, and A. J. Kellock, *J. Colloid Interface Sci.* 156, 240 (1993).
195. Ph. Espiard, A. Guyot, J. Perez, G. Vigier, and L. David, *Polymer* 36, 4397 (1995).
196. E. Bourgeat-Lami, Ph. Espiard, and A. Guyot, *Polymer* 36, 4385 (1995).
197. H. Xu, Z. Yang, H. Kong, S. Lu, and Y. Peng, *Polym. Mater. Sci. Eng.* 17, 99 (2001) (in Chinese).
198. J. Karger-Kocsis and J. Varga, *J. Appl. Polym. Sci.* 62, 291 (1996).
199. J. Karger-Kocsis, J. Varga, and G. W. Ehrenstein, *J. Appl. Polym. Sci.* 64, 2057 (1997).
200. J. Varga, in "Polypropylene: Structure, Blends and Composites" (J. Karger-Kocsis, Ed.), Vol. 1, p. 56. Chapman & Hall, London, 1995.
201. J. Varga and G. W. Ehrenstein, in "Polypropylene: An A-Z Reference" (J. Karger-Kocsis, Ed.), p. 51. Kluwer, Dordrecht, The Netherlands, 1999.
202. J. Liu, G. Tang, G. Tang, H. Zhou, and Q. Guo, *Chin. J. Appl. Chem.* 9, 26 (1992) (in Chinese).
203. C. Ye, J. Liu, and G. Tang, *Chin. Plast. Ind.* 22, 45 (1994) (in Chinese).
204. W. Wu and Z. Xu, *Acta Polym. Sinica* 1, 99 (2000) (in Chinese).
205. S. Ahmed and F. R. Jones, *J. Mater. Sci.* 25, 4933 (1990).
206. L. E. Nielsen and R. F. Landel, "Mechanical Properties of Polymers and Composites." Dekker, New York, 1994.
207. E. H. Kerner, *Proc. Phys. Soc.* B69, 808 (1956).
208. L. E. Nielsen, *J. Appl. Phys.* 4, 4626 (1970).

209. L. E. Nielsen, *J. Polym. Sci., Polym. Phys.* 17, 1897 (1979).
210. K. Ziegel, *J. Appl. Polym. Sci.* 17, 1119 (1973).
211. P. H. T. Vollenberg, J. W. de Haan, L. J. M. van de Ven, and D. Heikens, *Polymer* 30, 1663 (1989).
212. A. S. Ovchinskii, G. M. Gunyaev, T. G. Sorina, K. K. Borzunov, and I. S. Deev, *J. Adv. Mater.* 3, 304 (1996).
213. L. E. Nielsen, *J. Appl. Polym. Sci.* 10, 97 (1966).
214. L. Nicolais and M. Narkis, *Polym. Eng. Sci.* 11, 194 (1971).
215. L. Nicolais, E. Drioli, and R. F. Landel, *Polymer* 14, 21 (1973).
216. L. Nicolais and R. A. Mashelker, *J. Appl. Polym. Sci.* 20, 561 (1976).
217. J. Jancar, A. Dianselmo, and A. T. Dibenedetto, *Polym. Eng. Sci.* 32, 1394 (1992).
218. S. D. Sjoerdsma, *Polym. Commun.* 30, 106 (1989).
219. B. Pukanszky, B. Turcsanyi, and F. Tudos, in "Interfaces in Polymer, Ceramic and Metal Matrix Composites" (H. Ishida, Ed.), p. 467. Elsevier, New York, 1988.
220. E. Fekete, S. Z. Molnar, G.-M. Kim, G. H. Michler, and B. Pukanszky, *J. Macromol. Sci., Phys.* B38, 885 (1999).
221. C. Xiong, D. Wen, and Z. Pi, *Polym. Mater. Sci. Eng.* 4, 69 (1994) (in Chinese).
222. A. J. Kinloch and R. J. Young, "Fracture Behaviour of Polymers." Applied Science Publishers, London, 1983.
223. F. F. Lange and K. C. Radford, *J. Mater. Sci.* 6, 1197 (1971).
224. F. F. Lange, *J. Am. Ceram. Soc.* 54, 614 (1971).
225. Q. Fu and G. Wang, *Polym. Int.* 30, 309 (1993).
226. S. Wu, *Polymer* 26, 1855 (1985).
227. R. Zallen, "The Physics of Amorphous Solids." Wiley, New York, 1983.
228. K. Levon, A. Margolina, and A. Z. Patashinsky, *Macromolecules* 26, 4063 (1993).
229. M. Fujita, "Molecular Self-Assembly: Organic Versus Inorganic Approaches." Springer-Verlag, Berlin, 2000.
230. D. Y. Godovsky, *Adv. Polym. Sci.* 153, 163 (2000).
231. X. Chen, L. Huang, and D. Xu, *World Plast.* 13, 5 (1995). (in Chinese).
232. R. S. Sethi and M. T. Goosey, in "Special Polymers for Electronics and Optoelectronics" (J. A. Chilton and M. T. Goosey, Eds.), p. 1. Chapman & Hall, London, 1995.

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanoparticles as Drug Delivery Systems

Jörg Kreuter

Johann Wolfgang Goethe-Universität Frankfurt, Frankfurt, Germany

CONTENTS

1. INTRODUCTION

One of the primary objectives in the development of drug delivery systems is the controlled delivery of a drug to its site of action at an optimal rate [1] and in the most efficient way possible. Targeting a drug to its site of action would not only improve the therapeutic efficacy but also enable a reduction in total dose of the drug which must be administered to achieve a therapeutic response, thus minimizing unwanted toxic effects. This concept of site-specific or targeted delivery was already envisioned by Paul Ehrlich around the early 1900s when he proposed the idea of the concept of the "Zauberkugel" (English: magic bullet).

One possibility to reach this goal may be drug delivery by nanoparticles. Chiefly due to their small particle size below 1 μm, these drug delivery systems offer various advantages for many medical and veterinary applications [2, 3]. This particle size enables intravenous and intraarterial injection since particles of this size can easily traverse even the smallest blood capillaries with inner diameters of 3–8 μm [4]. A small particle size also minimizes possible irritant reactions at the injection site [5, 6].

2. DEFINITION OF NANOPARTICLES FOR PHARMACEUTICAL AND MEDICAL PURPOSES

Nanoparticles for pharmaceutical purposes are defined by the *Encyclopedia of Pharmaceutical Technology* [7] as solid colloidal particles ranging in size from 1 to 1000 nm (1 μm). They consist of macromolecular materials and can be used therapeutically as drug carriers, in which the active principle (drug or biologically active material) is dissolved, entrapped, or encapsulated, or to which the active principle is adsorbed or attached. This definition was already formulated earlier in a number of reviews [1, 3, 6, 8] and includes monolithic nanoparticles with drug dispersed in the polymer network or dissolved in the polymer in form of an amorphous solid solution (Fig. 1a) as well as capsules with a shell-like wall (Fig. 1b). It also includes for instance nanoparticulate adjuvants in vaccines. This definition does not include solid lipid nanoparticles which are produced from lipids [9–11]. Nevertheless, solid lipid nanoparticles are described in Section 4.9.

3. HISTORICAL

In the late 1960s and early 1970s Professor Peter Speiser at the ETH (Swiss Federal Institute of Technology) in Zürich realized Paul Ehrlich's idea and developed the first nanoparticles for drug delivery purposes [12–14]. These first nanoparticles were produced by emulsion polymerization of acrylamide crosslinked with *N*,*N*′-methylenebisacrylamide in hexane.

Independently, Zolle et al. [15] and Scheffel et al. [16] used another process, denaturation of albumin dissolved in water and emulsified in hot cottonseed oil, to produce nanoparticles. $^{99m}TcO_4^-$ was bound to the nanoparticles, and these particles were used for radioimaging of the lungs after intravenous injection.

Later the albumin nanoparticles were also used for drug delivery. Kramer [17] incorporated the anticancer drug mercaptopurine into these particles using the method of heat denaturation (Section 4.5).

The first successful drug targeting with nanoparticles was then performed by Widder and Senyei [18–22] by

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (161–180)

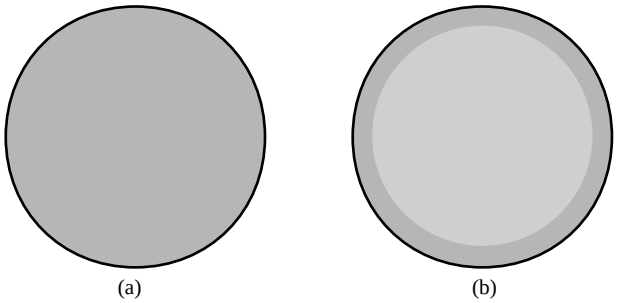


Figure 1. Schematic presentation of nanoparticles. (a) Monolythic nanoparticles with a continuous matrix. (b) Nanocapsules with a shell-type wall.

incorporation of magnetite particles into similar albumin nanoparticles and the employment of a magnetic field. These authors substituted mercaptopurine by the more efficient anticancer drug doxorubicin. Couvreur et al. [23–25] later could demonstrate that nonmagnetic nanoparticles also could accumulate in certain solid tumors and thus were able to substantially enhance the therapeutic efficacy of anticancer drugs such as doxorubicin or dactinomycin. This accumulation of the nanoparticles in these tumors probably was due to the enhanced permeability and retention effect (EPR effect) first described by Maeda and Matsumura [26] (see Section 6.1.1).

The latest important steps in the improvement of drug targeting with nanoparticles were the development of long circulating nanoparticles by the covalent linkage of polyethylene glycol (PEG) chains to poly(lactic-co-glycolic acid) (PLGA) [27] or to poly(alkyl cyanoacrylate) nanoparticles [28, 29] and the targeting of drugs to the brain across the blood–brain barrier [30].

4. PREPARATION METHODS FOR NANOPARTICLES

Several methods exist for the manufacture of nanoparticles. Table 1 lists the most important manufacturing methods. A very comprehensive compilation of preparation methods is given in the book *Colloidal Drug Delivery Systems* [6]. Only a few new considerably different manufacturing methods have been developed since the publication of this book; these include the above-mentioned methods to produce particles with covalently attached PEG chains and methods to attach linkers for drugs or targeting moieties. Following, the most frequently used preparations are briefly described.

4.1. Emulsion Polymerization

Emulsion polymerization in a continuous aqueous phase [33–93] is one of the most frequently used methods to produce nanoparticles. This method has a number of advantages. It represents a rapid production method that can rather easily be scaled up [7, 31]. The most important advantage, however, is that no organic solvents at all or well tolerated solvents like ethanol [64] are employed during this procedure.

To produce nanoparticles by emulsion polymerization a sparsely soluble monomer is dissolved in the aqueous phase. Additional monomer above its solubility may be added. This additional monomer can be stabilized by addition of emulsifiers (hence the name "emulsion polymerization") and then is present in emulsifier micelles or in larger emulsion droplets. However, the presence of an emulsifier is not necessary and does not influence the kinetics of this process [32–35]. The polymerization is initiated by radicals. These may be generated by the decay of initiator molecules such as potassium peroxodisulfate [35], which decays into two radicals upon heating, or by high energy radiation, such as γ-irradiation [37], ultraviolet, or strong visible light. The particle size as well as the molecular weights increase with increasing monomer concentration and decreasing initiator concentration [35, 36]. The polymerization is terminated by collision of two radicals or by two oppositely charged ions. The radicals or ions can be micro- or macroradicals and small ions or polymer ions, respectively. The termination can occur in the surrounding phase before phase separation or within the particles when two unterminated growing oligomer or polymer molecules meet [34, 38].

The heterogeneity in termination is one of the reasons polymerization of the poly(alkyl cyanoacrylates) is a rather complex procedure [54–57]. Poly(alkyl cyanoacrylates) have the advantage that they belong to the most rapidly biodegradable polymers. The polymerization of the poly(alkyl cyanoacrylates) is initiated by bases, specifically by the OH^- ions present in water. Actually, the pH has to be reduced below pH 3 in order to form nanoparticles. Above this pH the cyanoacrylates polymerize to large lumps. The pH may have to be even further reduced if basic components, for instance basic drugs, are present in the polymerization medium. On the other hand the water also generates the terminating species (i.e., H^+). For this reason the molecular weights of the poly(alkyl cyanoacrylates) in the nanoparticles are rather low [57] for instance compared to poly(methyl methacrylate) nanoparticles [36], and the influence of the monomer concentration on particle size is relatively small [54–56]. An increasing pH generally reduces the particle size, but again the situation is rather complex, and the pH influence is not very pronounced. Other components that are present in the polymerization mixtures such as drugs or stabilizers may have a much greater influence on the particle size and on the molecular weight [94, 95]. The presence of stabilizers tends to increase the molecular weights. The stabilizers, especially macromolecular stabilizers, which are present during polymerization are to a certain degree intermingled with the polymer network, most likely on the particle surface, but are not covalently bound to the polymer [96].

Nanoparticles also may be produced in a continuous organic phase. Solvents such as hexane, cyclohexane, isooctane, chloroform, and others were used. Polyacrylamide nanoparticles thus produced were among the first nanoparticles [12–14]. Later this process also was adapted for the production of poly(alkyl cyanoacrylate) nanoparticles [97–99]. However, the employment of organic solvents creates toxicological and environmental problems. For this reason, these processes are less useful.

Table 1. Preparation methods.

Method	Polymer	Ref.
Emulsion polymerization in a continuous aqueous phase	Poly(methyl methacrylate)	[33–43]
	Poly(methyl methacrylate)-copolymers	[44–53]
	Poly(butyl cyanoacrylate)	[30, 54–69]
	Poly(isobutyl cyanoacrylate)	[23–25, 70–81]
	Poly(hexyl cyanoacrylate)	[82–85]
	Poly(isohexyl cyanoacrylate)	[74–78, 81, 86–93]
Emulsion polymerization in a continuous organic phase	Poly acrylamide	[12–14]
	Poly(alkyl cyanoacrylates)	[97–99]
Interfacial polymerization	Poly(isobutyl cyanoacrylate)	[101–104, 107–112]
	Poly(isohexyl cyanoacrylate)	[105–107]
Solvent evaporation	Poly(lactic acid) and Poly(lactic acid-co-glycolic acid)	[115–143]
	Poly(ε-caprolactone)	[119, 128–130]
	Poly(ß-hydroxybutyrate)	[137, 138]
Solvent deposition or nanoprecipitation	Poly(lactic acid) and Poly(lactic acid-co-glycolic acid)	[145–148, 151, 154–157]
	Poly(ε-caprolactone)	[148–150]
	Eudragit®	[148, 154]
	Poly(ß-malic acid-cobenzylmaleate)	[152]
	Cellulose acetate phthalate	[153]
Preparation of nanoparticles in an oil emulsion	Albumin	[158–171]
	Gelatin	[172, 173]
	Chitosan	[174]
Desolvation of macromolecules	Gelatin	[178, 181, 184, 192–195]
	Albumin	[178, 182, 183, 187–191, 192–197]
	Cellulose	[186]
PEGylated nanoparticles	Poly(lactic acid) and Poly(lactic acid-co-glycolic acid)	[27, 200–205]
	Poly(alkyl cyanoacrylates)	[28, 29, 208–212]
	Albumin	[213–215]
Other methods	Polyglutaraldehyde	[216–218]
	Starch	[221]
	Polyacryl dextran	[222]
	Polyacryl starch	[223–226]
	Alginate	[227, 228]
	Chitosan	[229–231]
	Poly(ethylene imine)–fatty acid complexes	[232]
	Cholesterol–pullulan copolymers	[233, 234]

Emulsion polymerization is thermodynamically controlled. For this reason the particle size distribution generally is rather small.

4.2. Interfacial Polymerization

Interfacial polymerization occurs between two partners of a polymerization reaction that are soluble in two different immiscible solvents. The classical interfacial polymerization partners are hexamethylene amine, which is soluble in water, and a bifunctional acid chloride, for instance sebacyl chloride dissolved in chloroform or another organic solvent. The two bifunctional components react at the interface between the two liquids under formation of HCl yielding a polymer film, in the case of the above partners the producing Nylon® 6–10. This process was one of the first used to make microcapsules [100].

In order to make nanocapsules this process has to be modified. Significant amounts of a water soluble organic solvent—mainly ethanol is used—have to be present in the organic phase to help the dispersion of the organic phase in the aqueous medium. Poly(alkyl butyl cyanoacrylate) nanocapsules thus are produced by dissolution of the cyanoacrylate monomers in a mixture of ethanol and Miglyol® 812 or 829, or benzyl alcohol [101–107] plus drug. This mixture then is added slowly through a thin tube to an aqueous solution of a surfactant leading to the formation of a nanocapsule with a wall. Wall formation was demonstrated by electron microscopy [102, 105]. It is conceivable that this wall forms by polymerization of the cyanoacrylate

monomer present in the organic microemulsion droplets at the interface to the aqueous medium upon contact with the OH^- ions present in the surrounding water.

Chouinard et al. [107], however, argue that the polymer is already formed in the organic phase due to the nucleophilic character of ethanol and is then precipitated at the interface after emulsification in the aqueous phase similarly to the solvent deposition process described in Section 4.4.

Interfacial polymerization also can be carried out in an outer organic phase yielding nanocapsules with an aqueous core [111, 112]. In this case a 20% aqueous ethanol solution containing the drug is added to a surfactant containing (20%) Miglyol® 812 solution and finely dispersed by stirring. Isobutyl cyanoacrylate monomer is then added leading to the formation of nanocapsules within 4 h by interfacial polymerization. The nanoparticles are then separated from the organic phase by ultracentrifugation and resuspended in an aqueous medium.

4.3. Solvent Evaporation

Solvent evaporation is a classical method to produce microcapsules of the monolythic type [113, 114]. The manufacturing method can be modified to produce nanoparticles. In contrast to the *in-situ* polymerization processes mentioned in the previous sections, for solvent polymerization, a preformed polymer is used, usually poly(lactic acid) or poly(lactic-co-glycolic acid) [115–135, 139–144] or poly(ε-caprolactone) [119, 128–130]. The ratio between poly(lactic acid) and poly(glycolic acid) governs the biodegradation rate. The polymer is dissolved in an organic solvent. The drug also is contained in this solvent and may be present in the form of a solution or a fine dispersion, or it may be dissolved in water and then emulsified in the organic medium. As the next step, this organic solution or mixture is emulsified in an aqueous medium to form a O/W (oil in water) emulsion. After emulsification the organic liquid is evaporated under reduced pressure or by spray drying. This then leads to the formation of the solid polymer particles.

In order to form nanoparticles by the solvent evaporation process, compared to normal microcapsules much larger amounts of emulsifiers such as poloxamers [115], poloxamines [135], polysorbates [115, 132, 133], sodium lauryl sulfate [115], sodium dodecyl sulfate, poly(vinyl acetate) [118], poly(vinyl alcohol) [120, 124–126, 131], albumin [121–126], gelatin [117, 127], etc., plus the employment of high sheer homogenizers or extensive ultrasoniction are necessary in order to form the necessary small (nano)particle size. Since this process is kinetically controlled, the resulting particle size distribution is much broader than after emulsion polymerization. Besides the poly(lactates) other polymers have been used for this process including polyacrylates [136], ethyl cellulose, and poly(ß-hydroxybutyrate) [137, 138]. Some of the latter polymers, such as Eudragit® RL or RS and ethyl cellulose, even did not require the use of additional emulsifiers or surfactants due to their amphiphilic properties [136]. In some cases a water soluble organic solvent such as acetone may be added which aids in dispersion of the polymer. If these amounts are relatively minor [139–142] the formation of rather homogeneous monolithic particles prevails; higher amounts lead to formation of nanocapsules. This process is called solvent deposition.

The solvent evaporation process also may be used to encapsulate materials in the form of a double emulsion [143]. A very good review about solvent evaporation and related processes is given by Arshady [144].

4.4. Solvent Deposition or Nanoprecipitation

Solvent deposition or nanoprecipitation describes the formation of nanocapsules or nanoparticles by dissolution of poly(lactic acid) or a poly(lactic acid)-copolymer in a water miscible organic solvent, such as acetone [145–147] or benzyl alcohol [148]. Again the drug is dissolved in or admixed to the organic solution. This mixture is then poured into a surfactant-containing water phase under moderate stirring. Surfactants that can be used include poloxamers 118, 407, poloxamine 904, 908, and 1504, or stabilizing hydrocolloids such as poly(vinyl alcohol) or gelatin. The aqueous phase turns opalescent due to the formation of nanocapsules [145]. The acetone or a similar water miscible organic solvent is then removed similarly to solvent evaporation. Besides polylactic acid and copolymers, other polymers such as poly(ε-caprolactone) [148–150], Eudragit® [148] or another methyl methacrylic acid copolymer [151], poly(ß-malic acid-cobenzyl maleate) [152], or cellulose acetate phthalate [153] can be used, and by using acetonitrile as the organic solvent even no surfactants are required [151].

The process of precipitation–solvent evaporation can be modified by addition of the organic phase into a electrolyte rich aqueous solution. The electrolyte leads to the salting out of the organic materials and to the formation of nanoparticles after addition of water which results in the extraction of the water miscible organic solvent [153–157].

The solvent deposition method also may be called nanoprecipitation, solvent displacement, or the precipitation–solvent evaporation method.

4.5. Preparation of Nanoparticles in an Oil Emulsion

One of the first methods to produce nanoparticles was the emulsification of an aqueous solution containing a dissolved macromolecule such as albumin [15–22, 158–171], gelatin [172, 173], chitosan [174], or a similar macromolecule together with the drug in an oil. The emulsion is then homogenized with a high sheer homogenizer or by ultrasonication and poured into hot oil leading to the denaturation of the macromolecules and the formation of solid nanoparticles. Alternatively to heat denaturation the particles also may be hardened at much lower temperatures, even at room temperature, by cross-linking with an aldehyde [20, 172, 174] or by gelation [172]. After solidification of the nanoparticles the oil is removed by washing with volatile organic solvents.

The particle size and the size distribution of the nanoparticles are influenced by the type of oil, macromolecule concentration, aqueous-to-nonaqueous phase concentration, emulsification temperature, and surfactant used [6, 7, 160–166]. The time lapse between emulsification and hardening must be kept to a minimum to prevent an increase in particle size and distribution [6, 7]. Like nanoparticle formation by solvent evaporation, nanoparticle production in an

oil emulsion is kinetically controlled and yields a relatively broad particle size distribution.

References [175–177] represent comprehensive reviews about these preparation methods.

4.6. Desolvation of Macromolecules

Another method in which preferentially natural macromolecules are employed is the desolvation of macromolecules [178–197]. Macromolecules swell in a good solvent. The better the solvent and the greater the dilution, the more the molecules will be expanded. The process is reversible upon addition of a desolvating agent (nonsolvent), or changes in charge, temperature, and pH. This latter process is commonly known as coacervation [180] which results in the formation of a new phase called coacervate. Most commonly used microcapsules are produced by this process.

Shortly before phase separation occurs, the molecules are in a tightly packed, rolled-up conformation. If this structure can be maintained by addition of cross-linking agents, mainly aldehydes, or by heat denaturation [194], nanoparticles are formed. The process can be monitored by turbidity measurements, because due to the rolling-up the turbidity goes through a minimum. Upon addition of more desovating agent the formation of a coacervate starts which results in a strong increase in the turbidity (i.e., in the amount of scattered light). The method of desolvation of macromolecules is mainly used to produce albumin or gelatin nanoparticles, but also cellulose particles have been produced [186].

One of the problems that arises especially with gelatin is the heterogeneity in molecular weights of the employed macromolecules. This can lead to a wide range in particle sizes and to the formation of large aggregates. Therefore this manufacturing method was recently optimized by the invention of a two-step desolvation process [192–195]. In this case, after the first desolvation step, the lower molecular weight macromolecule fractions in the supernatant are removed after separation by a variety of methods and the sediment is redissolved again. After a second desolvation step the particles are hardened with an aldehyde.

Nanoparticles produced by this method can be surface-modified by introduction of sulfhydryl groups [193, 195, 197]. These groups can then be used to introduce certain linkers and ligands, for instance avidin [193, 195]. The avidin can then be used further to couple a diversity of other biotinylated ligands with therapeutic or targeting properties.

4.7. Pegylation of Poly(lactic Acid) and Poly(alkyl Cyanoacrylate) Nanoparticles

Nanoparticles like other colloidal carriers are rapidly taken up by cells of the reticulo-endothelial system (RES), also referred to as the mononuclear phagocytic system, and are then removed from the bloodstream [189, 199]. As a consequence, they do not remain long enough in the circulation to be able to reach the desired target site (unless of course the target site is the RES). One possibility to prolong the circulation time is the so-called "stealth" approach, the covalent attachment of PEG chains on the polymer. This process is called PEGylation (or pegylation) of the particles. The PEG chains are protruding from the nanoparticle surface into the aqueous environment yielding a hydrophilic shield around the particles and minimizing the adsorption of opsonins, compounds from the bloodstream that trigger the endocytotic uptake by macrophages and other cells of the RES.

The first nanoparticles that were pegylated were poly(lactide) nanoparticles [27, 200–205]. To obtain a coating that might prevent opsonization and subsequent recognition by the cells of the RES, PEG is attached to the particles using amphiphilic diblock copolymers, PEG-R, where R is the biodegradable polymer, that is, poly(lactic acid) or poly(lactic-co-glycolic acid) copolymer [27, 200, 201]. The PEG-R copolymers can be formed by the direct reaction of the terminal amino group from the monoamine monomethoxy PEG with the reactive carboxyl end group of the poly(lactic-co-glycolic acid). They can also be synthesized by the ring-opening polymerization at 114 °C of lactide and glycolide in the presence of monomethoxy PEG by the use of stannous octoate as a catalyst [27]. Instead of poly(lactic acid) or poly(lactic-co-glycolic acid) poly(ε-caprolactone) can be used [202–204]. The PEG-R polymers have amphiphilic properties where PEG due to its good water solubility represents the hydrophilic and R, being not water soluble but very soluble in organic phases, represents the hydrophobic part. For this reason, nanoparticles can easily be formed by the solvent evaporation method using methylene chloride or ethyl acetate [27] or by the solvent deposition or solvent displacement method using acetone [201]. Due to its hydrophilicity the PEG part of the copolymer enriches on the polymer–water interface and after removal of the organic by evaporation [27] or solvent displacement [201] solid particles are formed. Therefore, with pegylated polymers generally no additional emulsifier is required. The production can easily be optimized and particles from 100 [27, 200, 201] up to 820 nm can be prepared with good reproducibility [205]. Instead of diblock polymers, it is also possible to use PLA–PEG–PLA [PLA = poly(lactic acid)] using the salting-out process [206].

By a similar method sulfobutylated poly(vinyl alcohol)–graft–poly(lactide-co-glycolide) nanoparticles using acetone and the solvent displacement method were produced especially for the delivery of proteins [207].

Poly(alkyl cyanoacrylates) belong to most rapidly biodegrading polymers and may be even more rapidly degrading then the poly(lactides) [25, 126]. For this reason also pegylated poly(alkyl cyanoycralate) nanoparticles were produced [28, 29, 208–212]. One method uses emulsion polymerization of isobutyl cyanoacrylate in the presence of methylated PEG (Me_2PEG) as described in Section 4.1 [208]. The authors claim that the PEG is covalently linked to the poly(isobutyl cyanoacrylate) by this method. Later poly(methoxy-polyethyleneglycol cyanoacrylate-co-hexadecyl cyanoacrylate) nanoparticles were produced [28, 29, 209–212]. The methoxy-polyethyleneglycol cyanoacrylate-co-hexadecyl cyanoacrylate is amphiphilic like the pegylated poly(lactides). Consequently, as with the poly(lactides) nanoparticles can be produced by the solvent evaporation as well as by the nanoprecipitation methods [29].

In addition to these artificial polymers, PEG-modified albumin nanoparticles were prepared using albumin PEG

conjugates, poly(thioetheramido acid)-poly(ethylene glycol)-grafted human serum albumin and methoxy poly(ethylene glycol)-grafted human serum albumin [213–215]. The surface modified albumin nanoparticles released a model drug, rose bengal, in the presence of the enzyme trypsin much slower than similar unmodified albumin nanoparticles, suggesting the existence of a steric hydrophilic barrier by the PEG on the nanoparticle surface that made the digestion of the particles more difficult [215].

4.8. Nanoparticles Made by Other Methods

Nanoparticles also may be produced by a variety of other methods. These include polyglutaraldehyde nanoparticles produced by polycondensation techniques [216–220], carbohydrate nanoparticles made directly from starch [221] or from polyacryl dextran or polyacryl starch [222–226], alginate nanoparticles [227, 228], chitosan nanoparticles made by electrostatic interaction and complexing methods [229–231], nanoparticles consisting of poly(ethylene imine)–fatty acid complexes [232], and nanoparticles formed by self-aggregation of cholesterol–pullulan copolymers [233, 234]. This list, however, is not complete.

4.9. Solid Lipid Nanoparticles

Limitations in drug loading, a potential immunogenicity, and/or a too slow biodegradation rate are some problems that may occur with some of the polymers used for the preparation of nanoparticles. For this reason nanoparticles made from lipids were developed. These nanoparticles were called solid lipid nanoparticles (SLNs). Again, Speiser [235] was one of the pioneers in this field.

One of the most frequently used methods for the production of solid lipid nanoparticles is dispersion of warm O/W microemulsions in water [236–240]. The lipid or lipid mixture is melted, and water, drug, and surfactants with the same temperature of the melt are added. The right composition of this mixture this results in the formation of a thermodynamically stable microemulsion. This microemulsion is then poured into a cold (2–3 °C) aqueous phase which leads to the solidification of the lipids and the formation of solid lipid nanoparticles. Emulsifiers and dispersing agents used for this process include lecithin, bile salts, and butanol. Even the incorporation of a water phase into the lipid phase under formation of a W/O/W emulsion as an intermediate step is possible [240].

Another possibility for the production of solid lipid nanoparticles is the preparation by high pressure homogenization [9, 241–246]. This process again is kinetically controlled. Using high pressure homogenization the SLNs can be made in two ways, by hot homogenization [9, 241, 242] and by cold homogenization [243, 246]. For both techniques the drug is dissolved or solubilized in the liquid lipid which was heated to about 5 to 10 °C above its melting point. For the hot homogenization process this mixture is added to an aqueous surfactant solution of the same temperature and subjected to high pressure homogenization and then cooled down to room temperature. The lipid crystallizes and forms solid lipid nanoparticles. For cold homogenization the lipid melt is supercooled with liquid nitrogen or dry ice and milled in the supercooled stage. The resulting particles of a size of approximately 50 to 100 μm are then dispersed in a cold surfactant solution and also subjected to high pressure homogenization.

4.10. Inorganic Nanoparticles

Inorganic nanoparticles made of silica also were produced and their body distribution was investigated [247, 248]. Fumed silica with particle sizes around 16 nm (Aerosil® 200) is a frequently used excipient for oral, dermal, and rectal dosage forms. However, the use of silica nanoparticles for parenteral application is limited or not feasible due to the nonbiodegradability of this material. It has to be mentioned that colloidal aluminum hydroxide, phosphate, and oxide are contained as adjuvants in a number of intramuscularly or subcutaneously administered vaccines.

5. CHARACTERIZATION AND PROPERTIES OF NANOPARTICLES

The main important feature of nanoparticles is their size which enables them to traverse blood capillaries and may help to reduce irritant reactions at injection sites [4–6]. Some of the most important analytical methods for physicochemical characterization, drug loading, drug release, and *in vitro* interactions with blood components and cells are described in Sections 5.1–5.4 and Section 6.1. Many critical reviews have appeared about the characterization and properties of nanoparticles [6, 7, 33, 249–251].

5.1. Physicochemical Characterization

The important methods for the physicochemical characterization of nanoparticles are listed in Table 2. These methods mainly are similar to other materials or to larger particles.

Table 2. Physicochemical characterization methods for nanoparticles.

Parameter	Method	Ref.
Particle size	Photon correlation spectroscopy	[33, 249]
	Transmission electron microscopy	[33]
	Scanning electron microscopy	[33, 249]
	Scanning probe microscopy	[27]
	Fraunhofer diffraction	[249]
Molecular weight	Gel chromatography	[34, 36, 57, 90]
	Mass spectroscopy	
	Viscosimetry	[38]
	Analytical ultracentrifuge	
Crystallinity	X-ray diffraction	[33]
	Differential scanning calorimetry	
Polymer structure	Nuclear magnetic resonance	[96, 208]
Density	Helium compression pycnometry	[33]
Surface charge	Electrophoresis	[33, 249]
	Laser Doppler anemometry	[33]
Hydrophobicity	Hydrophobic interaction chromatography	[249]
	Contact angle measurement	[33, 252]

Some of the most frequently occurring problems are discussed in this section.

A special difficulty in nanoparticle characterization represents particle sizing. The method of choice for size and size distribution determination is photon correlation spectroscopy (PCS) based on laser light scattering because this method provides a rapid size analysis. However, this method is strongly influenced and skewed by larger particles present in the sample. Therefore, a confirmation of the particle size measurement by another method is advisable. Electron microscopy is especially optimal for this purpose. However, it has to be considered that for scanning electron microscopy a thin conducting layer has to be sputtered onto the sample. Since the thickness of this layer is not certain, errors in particle sizes may result [6, 7, 33]. Another excellent method for particle sizing is transmission electron microscopy. This method, however, is very laborious and enables only the analysis of small sample fractions. The same applies to scanning probe microscopy [27]. Transmission electron microscopy after freeze-fracturing [33] also enables the identification of nanocapsules. If this method is used for sizing it has to be taken into account that particles are not only fractured across the equator but also close to the poles of the particles which again skews the particle size analysis [35].

Four methods exist for the molecular weight determination of nanoparticles: gel permeation chromatography, mass spectroscopy, viscosimetry, and analytical ultracentrifuge. The results obtained with these methods may differ. Mass spectroscopy is mainly useful for smaller molecular weight nanoparticles such as the poly(alkyl cyanoacrylate) nanoparticles. For gel permeation chromatography [34, 35, 57, 90] and viscosimetry [38] the nanoparticles have to be soluble in a useful solvent. Moreover, often appropriate molecular weight standards for calibration are missing. Analytical ultracentrifuge analysis is very laborious and requires considerable expertise which is not available in most labs. None of these methods is very helpful for the molecular weight determination of nanoparticles made of cross-linked polymers which, therefore, is often impossible.

Nuclear magnetic resonance (NMR) was used to answer the question if components that were present during polymerization were covalently linked to the polymer. It was shown by this technique after repeated dissolution and precipitation steps that dextran present during polymerization of poly(alkyl cyanoacrylate) polymerization was not covalently linked to the cyanoacrylate polymer in significant amounts [96]. It was claimed, however, that methylated PEG was covalently linked to the cyanoacrylate after being present during polymerization [208]. However, this was not proven by repeated dissolution and precipitation. Hence it may be possible that the PEG chain are just entangled with the polycyanoacrylate chain network.

5.2. Drug Loading

The aspects of drug loading also are covered by a number of reviews [6, 7, 250]. Drugs may be bound to the nanoparticles by a number of ways: (a) incorporation into the interior of nanocapsules, (b) incorporation into monolithic nanoparticles in form of a solid solution or solid dispersion [253], (c) covalent binding [13, 14, 193, 195, 197], (d) electrostatic binding [48, 50, 52], and (e) surface adsorption [35, 254–256]. In most cases, the physicochemical properties of the drug determine the method of nanoparticle preparation and, hence, also the method of drug binding. Thermolabile drugs cannot be incorporated into nanoparticles produced involving heating. Hydrophilic drugs cannot be incorporated into hydrophobic nanoparticles without difficulties and vice versa. Consequently, in many cases drug loading is not sufficient.

Surface adsorption in general is governed by a Langmuir type of interaction [35, 254–256]. Surfactants generally reduce adsorption. In some cases, however, they improve adsorption [59, 257].

The precise determination of the drug content in certain cases may be a problem because of the colloidal nature of the carrier [7]. The problems involved have been reviewed in detail [258].

5.3. Drug Release

Drug release from nanoparticles generally is rather rapid, although also a very slow release has been observed [259, 260]. Again this topic was subject of number of reviews [6, 7, 250, 251, 258].

Release may occur by desorption of surface-bound drug, diffusion through the nanoparticle matrix or nanocapsule wall, matrix erosion, or a combined erosion–diffusion process [6, 7]. After injection into the bloodstream nanoparticles are rapidly opsonized by blood components [198, 261, 262] (see Section 6.1). This adsorptive blood component layer may strongly impede drug release.

In vitro drug release may be studied by a variety of methods [7]. Due to the small size of the nanoparticles the separation of the releasing particles from the rest of the sample represents a major problem. In principle this separation can be achieved in two ways, either by ultracentrifugation or by separation by a membrane, using for instance side-by-side diffusion cells with artificial or biological membranes, dialysis bags, ultrafiltration, and centrifugal ultrafiltration. All these techniques have the common disadvantage that a significant time lapse exists between immediate release from the particle and the time of sampling, because ultrafiltration, membrane diffusion, and ultracentrifugation are all time consuming processes. During this time lapse the release continues; therefore, it is not possible to obtain real time release rates. A very good alternative to study the release is the employment of substances that change their analytic profile, for instance color, by going from a hydrophobic into a hydrophilic environment when traversing from the particle surface into the release medium. Such a compound is the fluorescent marker Prodan®. This marker exhibits a hypsochromic shift from 424 nm in poly(lactate) to 512 nm in water [125]. Unfortunately, this method is compound specific and therefore can only be employed for model experiments.

Moreover, it has to be considered that in very many cases the *in vitro* release cannot be correlated with the *in vivo* situation [257].

5.4. Biodegradation

Biodegradation is a very important property of nanoparticles for most therapeutic uses, especially for parenteral administration (injection or implantation). Biodegradation has a profound influence on the drug release rate (see Section 5.3). In addition, with the possible exception of vaccines, the nanoparticle carrier material has to be eliminated rather rapidly to avoid an accumulation in the body.

Most natural polymers such as albumin, gelatin, or starch are biodegraded and eliminated rather rapidly due to enzymatic digestion [123–126, 263]. Poly(butyl cyanoacrylate) is also very rapidly biodegradable. It represents the most rapidly biodegrading artificial polymer. After 24 h already about 80% of the polymer is eliminated from body after intravenous injection of poly(butyl cyanoacrylate) nanoparticles [25]. The degradation and elimination of poly(hexyl cyanoacrylate) is considerably more slow [263–265]. Poly(alkyl cyanoacrylate) nanoparticles degrade in the body by enzymatic cleavage of the side chain ester bond, and the rate of this process depends on the ester side chain length [266–268]. The degradation was found to occur by surface erosion, because the particle size decreases immediately after incubation and does not show any lag period [269, 270]. Due to the low molecular weights of the poly(alkyl cyanoacrylates) [57, 94] the resulting water soluble polymer acid is the excreted rather rapidly from the body [6].

Poly(lactide) nanoparticles are rapidly degraded by enzymes, especially lipases [123–126]. Again this appears to be degradation on the surface comparable to the poly(alkyl cyanoacrylate) nanoparticles. Some coating materials, for instance poly(vinyl alcohol), largely prevent this degradation, since this coating material cannot be degraded by enzymes [124–126]. Therefore drug release from poly(lactide) nanoparticles coated with poly(vinyl alcohol) release occurs by diffusion only [125]. The degradation of the poly(lactide) nanoparticles from the surface contrasts the degradation process of the larger poly(lactide) microspheres ($\gg 1\ \mu$m) which degrade by hydrolysis and bulk erosion mainly from the center [271–273]. For this reason the particle size of microspheres changes very little during degradation for a considerable time.

Poly(methyl methacrylate) nanoparticles degrade very slowly over periods of years [274]. As a consequence, they are only useful as model particles or, more importantly, as adjuvants for vaccines, where a long persistence of the antigen and a long immune response are desired [39, 275, 276].

6. BODY DISTRIBUTION OF NANOPARTICLES

Although a small particle size also seems to be advantageous after localized administration [5] such as administration into the eye, injection or implantation into muscles, under the skin, joints, tumors, etc., the majority of nanoparticles are developed for intravenous injection or for oral administration. The small particle size enables a passage through even the smallest capillaries, and after oral administration they may enable a favorable gastrointestinal release for drugs with solubility problems or even an uptake through the intestinal mucosa [65, 82, 277–281]. Therefore, the investigation of the body distribution is of prime importance.

6.1. Body Distribution after Intravenous Injection

The main objective in the development of nanoparticles is drug targeting. As mentioned in the Introduction drug targeting enables high concentrations at the target site, that is, the desired place or action, whereas only little drug is distributed to the rest of the body. This can be done using the natural passive distribution pattern of the body for the carrier and therefore it is called passive targeting [7, 282]. However, the target sites that can be reached by this approach are very limited. Active targeting refers to the distribution achieved after the natural distribution pattern has been modified by alteration of the surface properties, binding of ligands, employment of magnetic nanoparticles, etc. For this reason, many more target sites can be reached or the drug concentrations at the target are much higher.

6.1.1. Passive Targeting

Nanoparticles are rapidly coated with blood components [6, 7, 198, 199, 249, 250] in the same manner as other colloidal carriers such as liposomes, microemulsions, erythrocyte ghosts, etc. This process is called opsonization. These opsonizing molecules trigger a rapid uptake of the colloidal drug carriers mainly by the sessile but also by the wandering macrophages of the RES [283]. As a consequence, shortly after intravenous injection of such carriers about 60 to 90% of the injected dose is distributed into the liver, 2–20% into the spleen, a varying amount into the lungs, and a smaller amount into the bone marrow [7, 198].

Some nanoparticles have the tendency to accumulate in certain solid tumors [25, 264]. This accumulation is due to the so-called EPR effect [26] that also can be observed with the other particulate colloidal carriers as well as with soluble macromolecules. The EPR effect can be attributed to two main factors [284–286]: First, the endothelium around the blood vessels in tumors often is very discontinuous and leaky and therefore allows the extravasation of larger particulates. Second, no lymphatic drainage exists in these tumors that normally would remove these particulates. However, in order to take good advantage of this effect, the particles have to stay long enough in the circulation to reach the tumor and must not be removed before by the macrophages. A prolonged circulation can be achieved by pegylation (see Section 4.7) or by coating with surfactants (see Section 6.1.2).

6.1.2. Active Targeting

Active targeting is much more difficult to achieve. A number of possibilities exist.

The first possibility that was employed was the alteration of the surface charge [287]. However, only a redistribution within the RES was obtainable.

Another approach is the incorporation of small (10–20 nm) magnetic Fe_3O_4 particles into the nanoparticle matrix [19, 19, 163–166, 174, 288–290] and the targeting by an external magnetic field. A localization in predesigned

areas of rat tails was possible with magnetic albumin nanoparticles, and very promising results were obtained with such nanoparticles in rats with a Yoshida tumor, a rat tail tumor [21, 22]. However, this tumor is located in a thin long extremity. It is much more difficult to focus the particles by a magnetic field deeper inside the body. For this reason, magnetic targeting so far has not found broad application.

An obvious drug targeting strategy would be the attachment of target specific ligands such as antibodies to the nanoparticles. This approach was attempted rather early [291–293]. One of the difficulties with this strategy is similar to the successful exploitation of the EPR effect, namely to obtain both a suitable coupling of the antibodies plus a sufficiently long enough circulation time to achieve high concentration at the target site. No successful targeting was achieved with nanoparticles with bound antibodies against an osteogenic sarcoma 788T in mice [294]. Instead the particles accumulated in the liver and the spleen to a higher degree than uncoated nanoparticles. Since antibodies are large proteins they may activate the removal mechanisms for foreign materials in the body. In addition, antibodies are very specific. For this reason the production of nanoparticles with bound antibodies on an industrial scale does not seem to be very feasible.

As previously stated, a prolongation of the circulation time of the nanoparticles is very desirable. One possibility to prolong this time is the so-called "stealth" approach, the covalent attachment of polyethylene glycol chains as described in Section 4.7. This largely decreases the adsorption of opsonins and very significantly reduces the uptake by macrophages [27–29, 200–215]. Another possibility to prolong the circulation time is the adsorption of surfactants on the surface of the nanoparticles [295–302]. Hydrophilc block copolymers, poloxamers (Pluronic®) and especially polxamines (Tetronic®), were able to considerably prolong the circulation time of model polystyrene and poly(methyl methacrylate) nanoparticles after intravenous injection to rats and rabbits. These surfactants adsorbed in the form of a monomolecular film coating the nanoparticles when added at a concentration of around 0.1%. Below this concentration, the particles behaved in their body distribution like uncoated particles whereas the body distribution changed very little above this concentration [302]. Other surfactants behaved differently and some yielded a higher accumulation in certain non-RES organs [299]. Polystyrene nanoparticles of a size of 60 nm coated with poloxamer 407, for instance, allowed a high accumulation in the bone marrow [303], and poly(butyl cyanoacrylate) nanoparticles coated with polysorbate 80 (Tween® 80) enabled the transport of significant amounts of a number of drugs across the blood–brain barrier (BBB) into the brain, both after intravenous injection [30, 66, 67, 304, 305]. Surface coating also plays a role in tumor targeting. Nanoparticles coated with different surfactants accumulated differently in different types of tumors [64, 306, 307]. It has to be mentioned that besides the surface coating, the nanoparticle material also plays a role in the body distribution obtained.

Both the prolongation of the circulation time as well as the enhancement in distribution into specific organs or certain tumors depend on the opsonins, complement [308], and other blood components which become adsorbed on the surface of the nanoparticles. The adsorption pattern of these blood components is governed by the surface properties of the nanoparticles [262, 309–311]. Depending on the surface properties, certain proteins are preferentially adsorbed after contact with the blood and some components will be enriched while others are reduced. Carriers with different protein adsorption patterns acquired following intravenous injection will interact with different tissue specific receptors or will be recognized by different macrophage subpopulations or different tumors. This adsorption pattern can be analyzed by two-dimensional polyacrylamide electrophoresis [262, 309–311]. Polysorbate 80-coated poly(butyl cyanoacrylate) nanoparticles for instance adsorb apolipoprotein E and/or B from the blood [312]. These apolipoproteins then seem to interact with the lipoprotein receptors on the brain blood capillary endothelial cells and thus are endocytosed by these cells leading to the above-mentioned delivery of drugs across the blood–brain barrier [68, 312].

6.2. Body Distribution after Subcutaneous and Intramuscular Injection

After subcutaneous and intramuscular injection, the nanoparticles remain at the injection site [274, 313]. After subcutaneous injection of ^{14}C-labelled poly(methyl methacrylate) nanoparticles, less than 1% of the dose distributed into the rest of the body.

6.3. Body Distribution after Oral Administration

The uptake of nanoparticles after oral administration has been studied in a number of investigations [278–281, 314–319]. A calculation of the total radioactivity excreted within 48 h after oral administration of ^{14}C-labelled poly(methyl methacrylate) nanoparticles via bile and urine amounted to over 10% of the administered dose [278]. Nevertheless, the radioactivity in individual organs was rather low. The highest concentrations were found in the liver, kidneys, muscles, and bone marrow [278, 280]. Up to 10% of the administered dose was found in the gastrointestinal walls. In general, gastrointestinal uptake of nanoparticles can occur by (a) a paracellular pathway, (b) intracellular uptake and transport via the epithelial cells lining the intestinal mucosa, and (c) lymphatic uptake via the M-cells and the Peyer's patches. Evidence exists for all three pathways, and the simultaneous existence of more than one pathway was reported [277]. Both surface properties [280, 318, 319] as well as the particle size [314] seem to play a role in the uptake pathway and in the amount of particles taken up. Concerning the influence of coadministered surfactants the results are contradictive; some studies suggest a reduced uptake [318, 319] whereas other investigators report a considerably enhanced uptake [280]. Again the different composition of the nanoparticle core between these studies might be responsible for the observed trend. Lymphatic uptake seems to be favored by hydrophobic particles [315–320].

7. APPLICATION OF NANOPARTICLES

Nanoparticles can be used as carriers for all sorts of drugs. The following section describes the most important applications.

7.1. Cancer

One of the most straightforward applications of nanoparticles is their employment in the targeting of anticancer drugs to solid tumors via the EPR effect (see Section 6.1.1). The accumulation of nanoparticles in certain solid tumors as well as the influence of the surface properties on this accumulation has been demonstrated in a number of studies [23, 25, 64, 306, 307].

The two cytostatic drugs of antibiotic origin, dactinomycin and doxorubicin, as well as another cytostatic, 5-fluorouracil, were among the first compounds that were coupled to nanoparticles and tested in animals using a variety of malignant experimental tumors [23, 24]. Dactinomycin bound to poly(alkyl cyanoacrylate) nanoparticles significantly reduced tumor growth. However, this beneficial effect was accompanied by a parallel increase in toxicity. The same was true with 5-fluorouracil nanoparticles [321]. In contrast, with doxorubicin the toxicity was significantly reduced compared to free drug [23], leading to an increased efficacy and safety [322]. In addition, doxorubicin bound to nanoparticles enabled an accumulation [25] and increased efficacy [323] in experimental lung and hepatic metastasis, which probably was due to the EPR effect. Moreover, in tissue cultures poly(alkyl cyanoacrylate) nanoparticles were able to overcome multidrug resistance [93, 324–326]. As stated above, nanoparticles have a tendency to be taken up especially by the macrophages of the RES [91]. This can explain why the drug efficacy after binding to nanoparticles was potentated by the activation of macrophages releasing cytotoxic factors such as NO, which resulted in increased cell death [327]. Because of the accumulation in the RES toxicological effects of the poly(alkyl cyanoacrylate) nanoparticles toward these organs were investigated [92, 328–332], and some toxicity was observed. However, this toxicity was largely reversible. Bone marrow toxicity was higher with poly(isohexyl cyanoacrylate) nanoparticles than with poly(isobutyl cyanoacrylate) nanoparticles [329]. This favorable toxicity led to the development of doxorubicin poly(isobutyl cyanoacrylate) and poly(isohexyl cyanoacrylate) nanoparticle preparations with bound doxorubicin for human application that were then subjected to a clinical trial [333, 334]. The doxorubicin-loaded nanoparticles exhibited a higher therapeutic index than free doxorubicin and also showed promise to overcome multidrug resistance. Although unexpected side effects such as fever, bone pain, or allergic reactions were noted, they were well tolerated and were all rapidly reversible [334].

An alternative to the binding of free doxorubicin is the coupling of this drug to dextran [335] or directly to poly(lactic-co-glycolic acid) [336, 337] and the incorporation of these conjugates in nanoparticles. An enhanced antitumoral effect also was observed with these preparations.

Other cytostatic drugs that have been bound to nanoparticles include mitoxantron [64, 306], cisplatin [338], and paclitacel [339, 340].

One strategy to increase drug delivery to a tumor is the attachment of antibodies as discussed in Section 6.1.2. Since antibodies are rather large and may encounter problems of immunogenicity after repeated injection, the attachment of receptor specific targeting moieties, such as vitamins and hormones, may be an alternative. The folic acid receptor, for instance, is such a target since it frequently is overexpressed on the surface of a variety of human cancer cells, especially ovarian carcinomas. For this reason, folic acid was coupled to poly(alkyl cyanoacrylate) nanoparticles and the specific interaction between these nanoparticles and the folate-binding protein was determined [341]. The analysis confirmed a specific binding of the nanoparticles to this protein which did not occur with nonconjugated nanoparticles.

Nanoparticles, therefore, are very promising for tumor targeting. Efficacy and toxicity depend strongly on the drug, the type of tumor [307], the surface properties, and the nanoparticle material used.

7.2. Antiinfective Agents

The treatment of intracellular infections often is difficult or impossible, since a number of drugs cannot reach certain target sites within these cells in which the microorganisms reside. Cells that frequently are infected and are inaccessible to a number of relevant antibiotics include phagocytic cells [6], such as the Kupffer cells in the liver, cells in the spleen, as well as circulating macrophages. Since nanoparticles are easily taken up by these cells, they are well suited for the delivery of antiinfective drugs to these cells. Possibilities for the targeted delivery of antibiotics by nanoparticles and liposomes have been reviewed recently [342].

The most promising results in treating such intracellular infections results were obtained with ampicillin bound to poly(isobutyl cyanoacrylate) and poly(isohexyl cyanoacrylate) nanoparticles [73, 86, 87]. Nanoparticle bound ampicillin was more effective against *Listeria monocytogenes* in mice at 20-fold lower concentrations than free ampicillin. The liver was free of bacteria 7 days after intravenous injection of the nanoparticle preparation. The bacterial count, however, increased again after this time. The nanoparticle preparation was similarly efficient against *Salmonella thyphimurium*. Although neither the liver nor the spleen was bacteria-free, all mice injected intravenously with the nanoparticles survived at an ampicillin dose that was 40-fold lower than required for free ampicillin.

Binding of primaquine to poly(isohexyl cyanoacrylate) nanoparticles also yielded a significant reduction in acute toxicity in mice and increased the life span of mice infected with *Plasmodium berghei* [343, 344].

A similar reduction of the toxicity was observed with dehydroemetine bound to poly(isohexyl cyanoacrylate) nanoparticles for the treatment of *Leishmania donovani* [345].

Other anti-infective agents that have been bound to nanoparticles include amikacin [62, 346], amphotericin B [142] ciprofloxacin [347], gentamycin [73], metronidazol [348], and the antiviral drug vidarabin [90]. Antiviral drugs against HIV (AIDS) are referred to in the next section.

7.3. AIDS

AIDS (acquired immune deficiency syndrome), an infection by the human immune deficiency viruses HIV 1 or 2, is mainly caused by infection of the T-helper cells and a subsequent breakdown of the immune system. While these cells die after infection, the virus can very well survive a number of other cells in the body, especially in macrophages. Since the macrophages easily take up nanoparticles, they are ideal delivery systems of anti-HIV drugs into these cells. The uptake of a variety of nanoparticles by HIV-infected and noninfected human macrophages was demonstrated *in vitro* [171, 187]. The uptake was even significantly higher in infected than in noninfected cells. No difference was observable between macrophages from infected patients and artificially *in vitro* infected cells [349]. Uptake also was influenced by the surface properties of the nanoparticles [187].

With the nucleoside analogs azidothymidine (AZT) and dideoxycytidine, the enhanced macrophage uptake only led to a small decrease in virus replication *in vitro* (macrophage cultures) because these drugs easily penetrate across cell membranes and get access to the cell interior [188]. In contrast, anti-HIV drugs such as the protease inhibitor saquinavir possess no or a very limited cell membrane permeability. Consequently this drug was already *in vitro* 10 to 20 times more efficient after binding to nanoparticles than with free drug or with a simple mixture of nanoparticles and free drug [83].

However, even with the drugs like AZT etc. that are able to penetrate cell membranes, binding to nanoparticles represents a major advantage. Cell cultures such as the macrophage cultures described above in most cases consist only of one cell type. The body, however, consists of numerous different cell types. Consequently, after intravenous injection to rats AZT bound to nanoparticles accumulated in the organs and cells of the RES that represented reservoirs for HIV, and, as a result, tissue concentrations in these organs were up to 20 times higher than with free drug [85, 86, 350]. Even after oral administration an enhanced uptake of nanoparticle bound drug AZT into these organs was observed [351].

7.4. Delivery to the Brain Across the Blood-Brain Barrier

The BBB represents an unsurmountable obstacle for a large number of drugs, including anticancer drugs, antibiotics, and a variety of central nervous system-active drugs, especially neuropeptides [352]. Nanoparticles coated with polysorbate 80 were shown to enable the delivery of a number of drugs across the BBB and to induce a significant pharmacological action [30, 66–68, 304, 305, 353–355]. Besides polysorbate 80, polysorbate 20, 40, and 60 also enabled this delivery, whereas with a variety of other surfactants no drug transport across the BBB was detectable [356]. The drugs that were delivered across the BBB with the polysorbate-coated nanoparticles include the hexapaptide dalargin [30, 66, 353] and another peptide, kytorphin [354], loperamide [67], tubocurarine [355], doxorubicin [304], and the NMDA-receptor antagonist MRZ 2/576 (8-chloro-4-hydroxy-1-oxo-1,2-dihydropyridazino[4,5-*b*]quinoline-5-oxide choline salt) [305]. Especially the results with the last two compounds are very promising: Doxorubicin seems to be very active against brain tumors but is not able to cross the BBB. After binding to polysorbate 80-coated nanoparticles drug concentrations in the brain were 60 times above the detection limit whereas all controls were below this limit [304]. No increase in doxorubicin toxicity appeared with these nanoparticles [69]. MRZ 2/576 is a potent but short acting anticonvulsant. Binding to the polysorbate-coated nanoparticles prolonged its action from 15 min up to almost 5 hours.

The mechanism for this drug delivery is not totally elucidated. *In vitro* an uptake of the polysorbate 80-coated nanoparticles but not of uncoated nanoparticles by endocytosis was observed in cultured human, bovine, rat, and mouse brain capillary endothelial cells [68, 357]. Only the polysorbate-coated nanoparticles adsorbed apolipoprotein E after incubation in plasma but not uncoated nanoparticles or those coated with the other surfactants. This apolipoprotein as well as apolipoprotein B but not other apolipoproteins also enabled a transport of significant amounts of the model hexapeptide drug dalargin across the BBB into the brain [312]. Both apolipoproteins are present in lipoprotein particles that are taken up by the lipoprotein receptors on the blood capillary endothelial cells of the brain. It is conceivable that the nanoparticles after injection into the bloodstream adsorb these apolipoproteins and thus mimic natural lipoprotein particles, leading to the uptake by these cells followed by drug delivery to the brain interior. Thus the polysorbate-coated nanoparticles would act as "Trojan horses." An alternative delivery mechanism, opening of the tight junctions between the endothelial cells, seems to be less likely since no major increase in the so-called inulin spaces was observed after intravenous injection [357], and no drug transport into the brain was observable with the other surfactants [54].

Transport of drugs across the BBB with nanoparticles was reviewed comprehensively in [352].

7.5. Oral Delivery

The basics for the employment of nanoparticles as oral drug delivery systems were discussed in Section 6.3 and are covered in a fundamental review [277]. Following, the oral delivery of a number of drugs is described.

7.5.1. Insulin

The oral delivery of the peptide insulin is a very important and yet unresolved issue in medicine since by this route of administration the presently necessary frequent injections can be avoided. Insulin is rapidly degraded in the gastrointestinal drug and therefore cannot be given in conventional oral dosage forms. Insulin can be encapsulated into poly(isobutyl cyanoacrylate) nanocapsules particles by the interfacial polymerization method described in Section 4.2 [101–103, 358–361]. The oral administration of these nanoparticles decreased glycemia in diabetic rats fasted overnight before dosing by 50–60% [103]. This effect lasted up to 20 days. Nonencapsulated insulin had no effect. In fed diabetic rats the nanoparticles yielded a 25% reduction in blood glucose levels only at high doses. Administration to different locations in the intestine exhibited a dependence of the glycemia on the administration

sites leading to 65% glycemia after administration into the ileum, 59% for administration into stomach, 52% into duodenum and jejunum, and 34% into colon [358]. Peak blood levels amounted to about 15 to 20% of the administered dose give parenterally [359]. *In vitro* investigations showed that the insulin is not chemically altered during nanoencapsulation [110]. The encapsulation into the nanoparticle process appears to protect the insulin against gastric degradations by enzymes. The interaction of nanoencapsulated insulin with insulin receptors *in vitro* was similar to that of native insulin regarding binding, negative cooperativity, and autophosphorylation [360].

7.5.2. Enhancement of the Bioavailability of Other Oral Drugs

The bioavailability and the drug action of a number of other drugs after oral administration also was improved considerably by binding to nanoparticles. The drugs showing an improved oral delivery with nanoparticles including vincamine [82], lipiodol [362], hydrocortisone [363], avarol [65], the anti-HIV drug azidothymidine, AZT [350, 351], the HIV 1 protease inhibitor CGP 57813 [156, 364], the peptide salmon calcitonin [365], and plasmid DNA, pCMV/ß-gal [366] as well as pCMVArah2 [367]. With indomethacin binding to nanoparticles did not result in an improved bioavailability after oral administration but rather in a reduction in side effects, and a prolonged action was observed [368, 369]. The most interesting results were those with avarol, where a nine-fold increase in bioavailability was obtained [65], and those with salmon calcitonin and plasmid DNA. Peptides as well as naked DNA are generally rapidly degraded in the gastrointestinal tract and they are poorly adsorbed.

Oral administration of salmon calcitonin bound to poly(*N*-isopropyl acrylamide) nanoparticles yielded a high hypocalcemic effect. The observed pronounced absorption enhancement by the nanoparticles probably resulted from both bioadhesion to the gastrointestinal mucosa and increased stability in the gastrointestinal tract [365]. Five days after a single oral dose of plasmid DNA, pCMV/ß-gal bound to fumaric acid-co-sebacic acid-polyanhydride nanoparticles was able to transfect cells in areas containing Peyer's patches [366]. Most of the transfected cells were located in the muscularis mucosae and adventita below these patches. Oral administration of pCMVArah2 complexed with chitosan to form nanoparticles resulted in transduced gene expression in the intestinal epithelium [367]. Mice receiving these nanoparticles produced secretory IgA and serum IgG2a. These results demonstrate that oral allergen-gene immunization with chitosan-DNA nanoparticles is effective in modulating murine anaphylactic responses and may be useful for treating food allergies.

The bioavailabiltiy after oral administration of drugs bound to nanoparticles may be further enhanced by the alteration of the surface properties [280]. Especially the coupling of tomato lectin seems to promising in promoting gastrointestinal uptake using nanoparticles [370].

7.6. Ocular Delivery

The efficacy of conventional ophthalmic dose forms is very poor because most of the instilled drug is drained away into the nose very rapidly, where it could lead to adverse effects. As a result, the bioavailability of eye drops is well below 5%. Therefore, an improvement in ocular delivery is very necessary. With gels and ointments the vision is strongly obstructed. Larger controlled or prolonged drug release dosage forms such as rods and microparticles can irritate the eye leading to an enhanced tear flow and a resulting increase in drug elimination by drainage. Due to their small size nanoparticles are much better suited for local delivery in the eye and a prolonged release. In addition, they possess slightly bioadhesive properties [371] which further can improve the ocular concentration of a variety of ophthalmic drugs. Nanoparticles with bound pilocarpine, one of the most important ophthalmic drugs, significantly increased the bioavailability of this drug, as determined by the intraocular aqueous humor drug concentration [60, 61]. This improvement led to a considerable prolongation in miosis time of over 30% and to an increase in intraocular pressure reduction from less than 5 to over 9 h; the latter action is representing the desired therapeutic effect. Besides the poly(alkyl cyanoacrylate) nanoparticles, cellulose acteate phathalate pseudolatex [372–374] and albumin [190, 191] nanoparticles were especially efficient in prolonging the action of nanoparticle-bound pilocarpine. The pharmacological action was further prolonged by the addition of bioadhesive polymers to the nanoparticle preparations [191].

Furthermore, the nanoparticles enabled a targeting to inflamed areas of the eye, away from inner eye segments where adverse effects could occur [189, 375]. This property makes them especially useful for the ophthalmic delivery of anti-inflamatory and anti-infective drugs.

Besides pilocarpine, other drugs that were bound to nanoparticles for ocular delivery include arecaidin propargylester [49], betaxolol [75], carteolol [376], metipranolol [377], hydrocortisone [189], amikacin [346], cyclosporine A [150], and progestereone [260]. The topic of ophthalmic application of nanoparticles also was covered in a review [378].

7.7. Delivery of DNA and Antisense Oligonucleotides

Another promising application of nanoparticles is their use as carriers for nucleic acids such as DNA and antisense oligonucleotides. A number of delivery issue requirements can be anticipated for genes and for oligonucleotides based upon their size, stability, the specific cell or tissue to be targeted, targeting concerns to specific tissues or cells, desired level of expression, and the potential for controlling that expression [379]. Some of these problems, especially issues of stability and targeting, may be overcome by the choice of an appropriate carrier. Viruses posses the ability to find their target cells and to permeate through cell membranes. However, large problems concerning immunological reactions were encountered with viral vectors. Therefore, nonviral delivery systems such as nanoparticles may be superior since they have the potential to be nonimmunogenic and stable *in vivo*.

The binding of these nucleic acid to the nanoparticle was achieved by three main strategies [380]:

1. The nucleic acids may be bound covalently either directly to the nanoparticles or better to hydrophobic linkers which than can be attached to the polymer surface by hydrophobic interaction [381]. Also binding via the avidin–biotin system is conceivable.
2. Cationic polymers such as cetyltrimethylammonium bromide (CTAB) [78] or DEAE-dextran [382] are attached to the nanoparticle surface. These polymers associate with the nanoparticle surface due to their amphiphilic character by intermingling with the nanoparticle polymer network when they are present during polymerization, or they can adsorb to the nanoparticle surface with their hydrophobic part. In both cases the cationic part of the molecules protrudes into the surrounding aqueous phase and the anionic nucleic acids bind to them by electrostatic interaction.
3. The nucleic acids may be incorporated into the nanoparticles. Although most nanoparticle polymer networks may be too tight to release the nucleic acids in a useful way [63], the incorporation into porous alginate nanosponges represents a promising approach [228].

The first experiments with plasmid DNA (pCAT-c and pUC-19 construct) bound to poly(butyl cyanoacrylate) nanoparticles [63] showed an excellent stability of the nanoparticle-bound DNA but no transfection in mice. The problems of plasmid encapsulation and release poly(lactide-co-glycolide) nanoparticles were studied in detail [383]. The most promising results for using nanoparticles were obtained with the chitosan [367] and the fumaric acid-co-sebacic acid-polyanhydride nanoparticles [366] which enabled the transfection of certain cells in the gastrointestinal region after oral administration as described in Section 7.5.2.

Silica nanoparticles also were used as carriers for plasmid DNA [384]. However, due to the nonbiodegradable and inorganic nature of these particles they may be less acceptable for administration to humans.

Antisense oligonucleotides are single-stranded short RNA or DNA fragments that are able to block the function of certain complementary sequences in mRNA or in genes. In the case of RNA this is achieved by mRNA Watson–Crick pairing thus inhibiting mRNA translation. This process is called translational arrest. DNA inactivation can occur by formation of a triple helix through Hoogsteen interaction. By employing the antisense strategy, certain genes or RNA function can be blocked thus preventing the replication of viruses or the proliferation of tumors by blocking oncogenes. The higher the number of bases the better the target specifity, but the more difficult it is to reach the target site.

The binding of antisense oligonucleotides to nanoparticles leads to a very pronounced protection against inactivation of the oligonucleotides by RNAses [51–53, 78, 79, 81, 88, 89, 382, 385–388]. The nanoparticles also improved the transport into the cell interior compared to free oligonucleotides and reduced or inhibited mRNA translation and gene expression. However, frequently oligonucleotides with a scrambled base sequence that was not complimentary to the sequence of the target nucleic acid and hence should not bind to the target also showed considerable inhibitory activity. This was observed for instance with an oligonucleotide designed to bind to the initiation site of the gag gene region of HIV 1 (GEM91) [387]: The encapsulation of this oligonucleotide into poly(lactic acid) nanoparticles and addition of these particles to HIV 1-infected CEM cells led to a strong inhibition of the viral production. Nanoparticle-entrapped random order gag oligonucleotides, however, were similarly effective.

Pharmacokinetic investigations in mice showed an increased accumulation of the nanoparticle-bound oligonucleotides in the liver reducing their distribution into the kidneys and bone marrow [79, 81]. The particles could partially protect the oligonucleotides against degradation in the plasma and in the liver for 5 min, whereas free oligonucleotide was totally degraded at the same time [79]. With alginate nanosponges 10-fold higher oligonucleotide concentrations were obtained in the lung than with poly(isobutyl cyanoacrylate) nanoparticles [81].

Poly(alkyl cyanoacrylate) nanoparticles were efficient against a tumor expressing the ras oncogen [89]. Oligonucleotides with bound antisense oligonucleotides directed to a point mutation in codon 12 of the Ha-ras gene were injected subcutaneously into nude mice. This point mutation leads to changes in cellular proliferation and induces the tumorigenic properties. The nanoparticle-bound oligonucleotides markedly inhibited Ha-ras-dependent tumor growth.

Whereas in most studies the oligonucleotides were attached to the nanoparticle surface and may be released prematurely, interfacial polymerization (see Section 4.2) allowed encapsulation into poly(isobutyl cyanoacrylate) nanocapsules [111, 112]. This may further improve the stability and the delivery of the oligonucleotides to the target cells.

7.8. Peptides and Proteins

The number of peptide and protein drugs is increasing steadily since they are becoming more and more assessable by using genetic engineering technology. However, their stability in body fluids, especially in the gastrointestinal tract, often is very low. In addition they often encounter difficulties in traversing membranes in the body. Peptides and proteins mostly bind rather well to nanoparticles, which can enable an improvement in the stability as well as in the transport across membranes and into specific cells.

The possibility of using nanoparticles to enable the oral delivery of the peptides insulin [103] and salmon calcitonin [365] and the transport of dalargin [30, 66] and kytorphin [354] was already shown and discussed in Sections 7.5.1 and 7.4, respectively.

The model protein bovine serum albumin (BSA) was incorporated into poly(lactic-co-glycolide) nanoparticles and their loading and release properties were investigated [389]. The release rate was fairly constant after an initial burst release and was enhanced by the incorporation of poloxamer 188, reaching approximately 100% with this preparation after 30 days. Other peptidic substances that were bound to nanoparticles include growth hormone releasing factor (hGRF) [77, 390], gangliosides [391], and granulocyte-colony stimulating factor (G-CSF) [392]. The stability of

these factors was increased and their release could be modulated. Poly(isohexyl cyanoacrylate) nanoparticles were able to maintain rather constant plasma levels of this drug for over 24 h after subcutaneous administration to rats. In contrast, with the same dose of free GRF plasma levels about 7 times higher appeared after 2 min that deceased rapidly, and no GRF was detectable any more after 100 min [390]. Mice experiments were conducted with rhG-CSF bound to poly(isohexyl cyanoacrylate) nanoparticles by addition after 8 h after start of the polymerization. However, no increase in short-term effects was observable after intravenous injection of the nanoparticle preparation compared to free (rhG-CSF) [392].

7.9. Vaccines

A large number of reviews have appeared on the subject of using nanoparticles as delivery systems and adjuvants for vaccines [393–397].

Poly(methyl methacrylate) nanoparticles were shown to be very good adjuvants for a number of antigens [40, 43, 275, 276, 398–402]. Antigens may be incorporated into these nanoparticles by polymerization in the presence of the antigen by gamma irradiation at low temperatures, or they may be adsorbed to previously polymerized particles [30, 35, 403]. Adsorption to previously polymerized particles also allows the employment of heat polymerization since most antigens are heat sensitive. The adjuvant effect depended on the particle size as well as on the hydrophobicity. The adjuvant effect increased with decreasing particle size [40, 398] and with increasing hydrophobicity [276, 400, 402]. Poly(methyl methacrylate) nanoparticles were superior to the conventional aluminum adjuvants which are the most frequently used adjuvants for human vaccination. The nanoparticles induced a good neutralizing antibody response against influenza [30] and good and long lasting protection [276, 399]. They were especially effective for subunit vaccines [275]. Poly(methyl methacrylate) nanoparticles not only yielded an excellent adjuvant effect but also increased the stability of the vaccine considerably [276].

Besides influenza [30, 275, 276, 399] and the model antigen BSA [40, 400], poly(methyl methacrylate) nanoparticles also were excellent adjuvants for HIV 1 and HIV 2 [43, 401]. A study involving a total of 24 different adjuvants demonstrated that the antibody response was dependent on the specific adjuvant used [402]. Different adjuvants led to large differences in antibody titers against the different antigenic components contained in the vaccines.

Poly(methyl methacrylate) nanoparticles are well tolerated by the body after subcutaneous injection. They are very slowly biodegradable [274] which is very beneficial because this enables a prolonged antigen presentation to the immune system of the body.

In addition, nanoparticles may be useful for oral vaccination. As mentioned in Section 7.5.2 oral administration of plasmid DNA (pCMVArah2) bound to nanoparticles produced secretory IgA and serum IgG2a [367]. Such nanoparticles may be useful for treating food allergies.

8. CONCLUSIONS

Nanoparticles represent very diverse colloidal drug delivery systems for a number of drugs, including nucleic acids and peptides. They also are promising adjuvants for vaccines. Nanoparticles enable the drug targeting to many solid tumors and the delivery of a number of drugs across the blood–brain barrier. They can significantly enhance the oral bioavailability and prolong the action of ophthalmic drugs.

GLOSSARY

Active targeting Targeting of particulate or soluble drug carriers to special target sites in the body after altering the surface properties, incorporation of magnetite particles, binding of ligands, etc., whereas passive targeting refers to the natural body distribution of these carriers without these alterations.

Drug loading Drug binding or amount of drug bound to nanoparticles.

Nanoprecipitation Formation of nanoparticles by precipitation of a water insoluble polymer dissolved in a water miscible organic solvent upon addition to water.

Solvent deposition Synonym to nanopreciptation.

REFERENCES

1. J. Kreuter, in "Microcapsules and Nanoparticles" (M. Donbrow, Ed.), p. 125. CRC Press, Boca Raton, FL, 1991.
2. J. J. Marty and R. C. Oppenheim, *Austral. J. Pharm. Sci.* 6, 65 (1977).
3. J. Kreuter, *Pharm. Acta Helv.* 58, 196 (1983).
4. G. Thews, E. Mutschler, and P. Vaupel, "Anatomie, Physiologie, Pathophysiologie des Menschen," p. 217. Wissenschaftliche Verlagsgesellschaft mbH, Stuttgart, 1999.
5. K. Little and J. Parkhouse, *Lancet* II, 857 (1962).
6. J. Kreuter, in "Colloidal Drug Delivery Systems" (J. Kreuter, Ed.), p. 219. Dekker, New York, 1994.
7. J. Kreuter, in "Encyclopedia of Pharmaceutical Technology" (J. Swarbrick and J. C. Boylan, Eds.), Vol. 10, p. 165. Dekker, New York, 1994.
8. J. Kreuter, in "Methods in Enzymology" (K. J. Widder and R. Green, Eds.), Vol. 112, p. 129. Academic Press, Orlando, FL, 1985.
9. C. Schwarz, W. Mehnert, J. S. Lucks, and R. H. Müller, *J. Controlled Rel.* 30, 83 (1994).
10. R. H. Müller, S. Maassen, C. Schwarz, and W. Mehnert, *J. Controlled Rel.* 47, 261 (1997).
11. A. Bargoni, R. Cavalli, O. Caputo, A. Fundaro, M. R. Gasco, and G. P. Zara, *Pharm. Res.* 15, 745 (1998).
12. G. Birrenbach and P. P. Speiser, *J. Pharm. Sci.* 65, 1763 (1976).
13. H. Kopf, R. K. Joshi, M. Soliva, and P. P. Speiser, *Pharm. Ind.* 38, 281 (1976).
14. H. Kopf, R. K. Joshi, M. Soliva, and P. P. Speiser, *Pharm. Ind.* 39, 993 (1977).
15. I. Zolle, F. Hosain, B. A. Rhodes, and H. N. Wagner, Jr., *J. Nucl. Med.* 11, 379 (1970).
16. U. Scheffel, B. A. Rhodes, T. K. Natarajan, and H. N. Wagner, Jr., *J. Nucl. Med.* 13, 498 (1972).
17. P. A. Kramer, *J. Pharm. Sci.* 63, 1646 (1974).
18. K. J. Widder, A. E. Senyei, and D. G. Scarpelli, *Proc. Soc. Exp. Biol. Med.* 58, 141 (1978).
19. A. E. Senyei, K. J. Widder, and Czerlinski, *J. Appl. Phys.* 49, 3578 (1978).

20. K. J. Widder, G. Flouret, and A. E. Senyei, *J. Pharm. Sci.* 68, 79 (1979).
21. K. J. Widder, R. M. Morris, G. A. Poore, D. P. Howard, and A. E. Senyei, *Eur. J. Cancer* 19, 135 (1983).
22. K. J. Widder, P. A. Marino, R. M. Morris, D. P. Howard, G. A. Poore, and A. E. Senyei, *Eur. J. Cancer* 19, 141 (1983).
23. P. Couvreur, L. Grislain, V. Lenaerts, F. Brasseur, P. Guiot, and A. Biernacki, in "Polymeric Nanoparticles and Microspheres" (P. Guiot and P. Couvreur, Eds.), p. 27. CRC Press, Boca Raton, FL, 1986.
24. F. Brasseur, P. Couvreur, B. Kante, L. Deckers-Passau, M. Roland, C. Deckers, and P. Speiser, *Eur. J. Cancer* 16, 1441 (1980).
25. L. Grislain, P. Couvreur, V. Lenaerts, M. Roland, D. Deprez-De Campenere, and P. Speiser, *Int. J. Pharm.* 15, 335 (1983).
26. Y. Matsumura and H. Maeda, *Cancer Res.* 46, 6387 (1986).
27. R. Gref, Y. Minamitake, M. T. Peracchia, V. Trubetskoy, V. Torchilin, and R. Langer, *Science* 263, 1600 (1994).
28. M. T. Peracchia, D. Desmaele, P. Couvreur, and J. d'Angelo, *Macromolecules* 30, 846 (1997).
29. M. T. Peracchia, C. Vauthier, D. Desmaele, A. Gulik, J.-C. Dedieu, M. Demoy, J. d'Angelo, and P. Couvreur, *Pharm. Res.* 15, 550 (1998).
30. J. Kreuter, R. N. Alyautdin, D. A. Kharkevich, and A. A. Ivanov, *Brain Res.* 674, 171 (1995).
31. J. Kreuter, in "Specialized Drug Delivery Systems" (P. Tyle, Ed.), p. 257. Dekker, New York, 1990.
32. R. M. Fitch, *Br. Polymer J.* 5, 467 (1973).
33. J. Kreuter, *Int. J. Pharm.* 14, 43 (1983).
34. J. Kreuter, *J. Polymer Sci., Polym. Lett. Ed.* 20, 543 (1982).
35. U. E. Berg, J. Kreuter, P. P. Speiser, and M. Soliva, *Pharm. Ind.* 48, 75 (1986).
36. V. Bentele, U. E. Berg, and J. Kreuter, *Int. J. Pharm.* 13, 109 (1983).
37. J. Kreuter and H. J. Zehnder, *Radiation Effects* 35, 161 (1978).
38. C. Marburger and J. Kreuter, *Progr. Colloid Polym. Sci.* 109, 270 (1998).
39. J. Kreuter, J. and P. P. Speiser, *Infect. Immunity* 13, 204 (1976).
40. J. Kreuter, U. E. Berg, E. Liehl, M. Soliva, and P. P. Speiser, *Vaccine* 4, 125 (1986).
41. A. Malaiya and S. P. Vyas, *J. Microencapsul.* 5, 243 (1988).
42. S. P. Vyas and A. Malaiya, *J. Microencapsul.* 6, 431 (1989).
43. F. Stieneker, J. Kreuter, and J. Löwer, *AIDS* 5, 431 (1991).
44. A. Rolland, *Int. J. Pharm.* 42, 145 (1988).
45. A. Rolland, J.-M. Bégué, R. Le Verge, and A. Guillouzo, *Int. J. Pharm.* 53, 67 (1989).
46. A. Rolland, B. Collet, R. Le Verge, and L. Toujas , *J. Pharm. Sci.* 78, 481 (1989).
47. K. Langer, C. Marburger, A. Berthold, J. Kreuter, and F. Stieneker, *Int. J. Pharm.* 137, 67 (1996).
48. K. Langer, F. Stieneker, G. Lambrecht, E. Mutschler, and J. Kreuter, *Int. J. Pharm.* 158, 211 (1997).
49. K. Langer, E. Mutschler, G. Lambrecht, D. Mayer, G. Troschau, F. Stieneker, and J. Kreuter, *Int. J. Pharm.* 158, 219 (1997).
50. F. Hoffmann, J. Cinatl, Jr., H. Kabickova, J. Cinatl, J. Kreuter, and F. Stieneker, *Int. J. Pharm.* 157, 189 (1997).
51. H.-P. Zobel, A. Zimmer, S. Atmaca-Abdel Aziz, M. Gilbert, D. Werner, C. R. Noe, J. Kreuter, und F. Stieneker, *Eur. J. Pharm. Biopharm.* 47, 203 (1999).
52. H.-P. Zobel, F. Stieneker, S. Atmaca-Abdel Aziz, M. Gilbert, M., D. Werner, C. R. Noe, J. Kreuter, J. and A. Zimmer, *Eur. J. Pharm. Biopharm.* 48, 1 (1999).
53. H. P. Zobel, M. Junghans, V. Maienschein, D. Werner, M. Gilbert, H. Zimmermann, C. Noe, J. Kreuter, and A. Zimmer, *Eur. J. Pharm. Biopharm.* 49, 203 (2000).
54. S. J. Douglas, L. Illum, S. S. Davis, and J. Kreuter, *J. Colloid Interf. Sci.* 101, 149 (1984).
55. S. J. Douglas, L. Illum, and S. S. Davis, *J. Colloid Interf. Sci.* 103, 154 (1985).
56. L. Illum, M. A. Khan, E. Mak, and S. S. Davis, *Int. J. Pharm.* 30, 17 (1986).
57. M. A. El-Egakey, V. Bentele, and J. Kreuter, *Int. J. Pharm.* 13, 349 (1983).
58. J. Kreuter and H. R. Hartmann, *Oncology* 40, 363 (1983).
59. T. Harmia, P. P. Speiser, and J. Kreuter, *Int. J. Pharm.* 33, 45 (1986).
60. R. Diepold, J. Kreuter, J. Himber, R. Gurny, V. H. L. Lee, J. R. Robinson, M. F. Saettone, and O. E. Schnaudigel, *Graefe's Arch. Clin. Exp. Ophthalmol.* 227, 188 (1989).
61. A. Zimmer, E. Mutschler, G. Lambrecht, D. Mayer, and J. Kreuter, *Pharm. Res.* 11, 1435 (1994).
62. M. J. Alonso, C. Losa, P. Calvo, and J. L. Villa-Jato, *Int. J. Pharm.* 68, 69 (1991).
63. W. M. Bertling, M. Gareis, V. Paspaleeva, A. Zimmer, J. Kreuter, E. Nürnberg, and P. Harrer, *Biotech. Appl. Biochem.* 13, 390 (1991).
64. P. Beck, J. Kreuter, R. Reszka, and I. Fichtner, *J. Microencapsul.* 10, 101 (1993).
65. P. H. Beck, J. Kreuter, W. E. G. Müller, and W. Schatton, *Eur. J. Pharm. Biopharm.* 40, 134 (1994).
66. R. Alyautdin, D. Gothier, V. Petrov, D. Kharkevich, and J. Kreuter, *Eur. J. Pharm. Biopharm.* 41, 44 (1995).
67. R. N. Alyautdin, V. E. Petrov, K. Langer, A. Berthold, D. A. Kharkevich, and J. Kreuter, *Pharm. Res.* 14, 325 (1997).
68. P. Ramge, R. E. Unger, J. B. Oltrogge, D. Zenker, D. Begley, J. Kreuter, and H. von Briesen, *Eur. J. Neurosci.* 12, 1931 (2000).
69. S. E. Gelperina, A. S. Khalansky, I. N. Skidan, Z. S. Smirnova, A. I. Bobruskin, S. E. Severin, B. Turowski, F. E. Zanella, and J. Kreuter, *Toxicol. Lett.* 126, 131 (2002).
70. P. Couvreur, B. Kante, M. Roland, P. Guiot, P. Bauduin, and P. Speiser, *J. Pharm. Pharmacol.* 31, 331 (1979).
71. P. Couvreur, B. Kante, L. Grislain, M. Roland, and P. P. Speiser, *J. Pharm. Sci.* 71, 790 (1982).
72. P. S. Bajwa, P. Couvreur, and P. A. Volz, *FEMS Microbiol. Lett.* 44, 412 (1987).
73. S. Henry-Michelland, M. J. Alonso, A. Andremont, P. Maincent, J. Saucières, and P. Couvreur, *Int. J. Pharm.* 35, 121 (1987).
74. C. Kubiak, L. Manail, and P. Couvreur, *Int. J. Pharm.* 41, 181 (1988).
75. L. Marchal-Heussler, P. Maincent, M. Hoffman, J. Spittler, and P. Couvreur, *Int. J. Pharm.* 58, 115 (1990).
76. B. Seijo, E. Fattal, L. Roblot-Treupel, and P. Couvreur, *Int. J. Pharm.* 62, 1 (1990).
77. J. L. Grangier, M. Puygrenier, J. C. Gautier, and P. Couvreur, *J. Controlled Rel.* 15, 3 (1991).
78. C. Chavany, T. Le Doan, P. Couvreur, F. Puisieux, and C. Hélène, *Pharm. Res.* 9, 441 (1992).
79. Y. Nakada, E. Fattal, M. Foulquier, and P. Couvreur, *Pharm. Res.* 13, 38 (1996).
80. M. Guzman, J. Molpeceres, F. Garcia, M.-R. Aberturas, and M. Rodriguez, *J. Pharm. Sci.* 82, 498 (1993).
81. E. Fattal, C. Vauthier, I. Aynie, Y. Nakada, G. Lambert, C. Malvy, and P. Couvreur, *J. Controlled Rel.* 53, 137 (1998).
82. P. Maincent, R. Le Verge, P. Sado, P. Couvreur, and J.-P. Devissaguet, *J. Pharm. Sci.* 75, 955 (1986).
83. A. R. Bender, H. von Briesen, J. Kreuter, I. B. Duncan, and H. Rübsamen-Waigmann, *Antimicrob. Agents Chemother.* 40, 1467 (1996).
84. R. Löbenberg and J. Kreuter, *AIDS Res. Human Retrovir.* 12, 1709 (1996).
85. R. Löbenberg, L. Araujo, H. von Briesen, E. Rodgers, and J. Kreuter, *J. Controlled Rel.* 50, 21 (1998).

86. M. Youssef, E. Fattal, M.-J. Alonso, L. Roblot-Treupel, J. Sauzières, C. Tancrède, A. Omnès, P. Couvreur, and A. Andremont, *Antimicrob. Agents Chemother.* 32, 1204 (1988).
87. E. Fattal, M. Youssef, P. Couvreur, and A. Andremont, *Antimicrob. Agents Chemother.* 33, 1504 (1989).
88. C. Chavany, T. Saison-Behomoaras, T. Le Doan, F. Puisieux, P. Couvreur, and C. Hélène, *Pharm. Res.* 11, 1370 (1994).
89. G. Schwab, C. Chavany, I. Duroux, G. Goubin, J. Lebeau, C. Hélène, and T. Behmoaras, *Proc. Natl. Acad. Sci. USA* 91, 10460 (1994).
90. V. Guise, J.-Y. Drouin, J. Benoit, J. Mahuteau, P. Drumont, and P. Couvreur, *Pharm. Res.* 7, 736 (1990).
91. C. Verdun, F. Brasseur, H. Vranckx, P. Couvreur, and M. Roland, *Cancer Chemother. Pharmacol.* 26, 13 (1990).
92. N. Chiannilkulchai, N. Ammoury, B. Caillou, J.-P. Devissaguet, and P. Couvreur, *Cancer Chemother. Pharmacol.* 26, 122 (1990).
93. C. Cuvier, L. Roblot-Treupel, J. M. Millot, S. Chevillard, M. Manfait, and P. Couvreur, *Biochem. Pharmacol.* 44, 509 (1992).
94. S. J. Douglas, S. S. Davis, and S. R. Holding, *Br. Polym. J.* 17, 339 (1985).
95. L. Vansnick, P. Couvreur, D. Christiaens-Leyh, and M. Roland, *Pharm. Res.* 1, 36 (1985).
96. S. Pirker, J. Kruse, Ch. Noe, K. Langer, A. Zimmer, and J. Kreuter, *Int. J. Pharm.* 128, 189 (1996).
97. H.-J. Krause, A. Schwarz, and P. Rohdewald, *Drug Devel. Ind. Pharm.* 12, 527 (1986).
98. M. S. El-Samaligy, P. Rohdewald, and H. A. Mahmoud, *J. Pharm. Pharmacol.* 38, 216 (1986).
99. R. Campignano, M. R. Gasco, and S. Morel, *Pharm. Acta Helv.* 66, 47 (1991).
100. T. M. S. Chang, *Science* 146, 524 (1964).
101. N. Al Khouri Fallouh, L. Roblot-Treupel, H. Fessi, J. P. Devissaguet, and F. Puisieux, *Int. J. Pharm.* 28, 125 (1986).
102. J. M. Rollot, P. Couvreur, L. Roblot-Treupel, and F. Puisieux, *J. Pharm. Sci.* 75, 361 (1986).
103. C. Damge, C. Michel, M. Aprahamian, and P. Couvreur, *Diabetes* 37, 246 (1988).
104. B. Hubert, J. Atkinson, M. Guerret, M. Hoffmann, J. P. Devissaguet, and P. Maincent, *Pharm. Res.* 8, 734 (1991).
105. F. Chouinard, F. W. K. Kan, J.-C. Leroux, C. Foucher, and V. Lenaerts, *Int. J. Pharm.* 72, 211 (1991).
106. S. Bonduelle, C. Foucher, J.-C. Leroux, F. Chouinard, C. Cadieux, and V. Lenaerts, *J. Microencapsul.* 9, 173 (1992).
107. F. Chouinard, S. Buczkowski, and V. Lenaerts, *Pharm. Res.* 11, 869 (1994).
108. P. Sai, C. Damge, A. S. Rivereau, A. Hoeltzel, and A. Gouin, *J. Autoimmun.* 9, 713 (1996).
109. C. Damge, J. Vonderscher, P. Marbach, and M. Pinget, *J. Pharm. Pharmacol.* 49, 949 (1997).
110. M. Aboubakar, F. Puisieux, P. Couvreur, and C. Vauthier, *Int. J. Pharm.* 183, 63 (1999).
111. G. Lambert, E. Fattal, H. Pinto-Alphandary, A. Gulik, and P. Couvreur, *Pharm. Res.* 17, 707 (2000).
112. G. Lambert, E. Fattal, H. Pinto-Alphandary, A. Gulik, and P. Couvreur, *Int. J. Pharm.* 214, 13 (2001).
113. P. B. Deasy, "Microencapsulation and Related Processes," p. 85. Dekker, New York, 1984.
114. E. Mathiowitz and R. Langer, in "Microcapsules and Nanoparticles in Medicin and Pharmacy" (M. Donbrow, Ed.), p. 105. CRC Press, Boca Raton, 1991.
115. R. Gurny, N. A. Peppas, D. D. Harrington, and G. S. Banker, *Drug Devl. Ind. Pharm.* 7, 1 (1981).
116. T. R. Tice and R. M. Gilley, *J. Controlled Rel.* 2, 343 (1985).
117. H. Krause, A. Schwarz, and P. Rohdewald, *Int. J. Pharm.* 17, 145 (1985).
118. Y. Tabata and Y. Ikada, *Pharm. Res.* 6, 296 (1989).
119. M. D. Coffin and J. W. Mc Ginity, *Pharm. Res.* 9, 200 (1992).
120. H. Jeffery, S. S. Davis, and D. T. O'Hagan, *Int. J. Pharm.* 77, 169 (1991).
121. D. V. Bazile, C. Ropert, P. Huve, T. Verrachia, M. Marland, A. Frydman, M. Veillard, and G. Spenlehauer, *Biomaterials* 13, 1093 (1992).
122. T. Verrachia, P. Huve, D. Bazile, M. Veillard, G. Spenlehauer, and P. Couveur, *J. Biomed. Mater. Res.* 27, 1019 (1993).
123. F. B. Landry, D. V. Bazile, G. Spenlehauer, M. Veillard, and J. Kreuter, *Biomaterials* 17, 715 (1996).
124. F. B. Landry, D. V. Bazile, G. Spenlehauer, M. Veillard, and J. Kreuter, *S.T.P. Pharma Sci.* 6, 195 (1996).
125. F. B. Landry, D. V. Bazile, G. Spenlehauer, M. Veillard, and J. Kreuter, *J. Controlled Rel.* 44, 227 (1997).
126. F. B. Landry, D. V. Bazile, G. Spenlehauer, M. Veillard, and J. Kreuter, *J. Drug Target.* 6, 293 (1998).
127. E. A. Poyner, H. O. Alpar, A. J. Almeida, M. D. Gamble, and M. R. W. Brown, *J. Controlled Rel.* 35, 41 (1995).
128. L. Marchand-Heusler, H. Fessi, J. P. Devissaguet, M. Hoffman, and P. Maincent, *S.T.P. Pharma Sci.* 2, 98 (1992).
129. M. C. Varela, M. Guzman, J. Molpeceres, M. del Rosario-Aberturas, D. Rodriguez-Puyol, and M. Rodriguez-Puyol, *Eur. J. Pharm. Sci.* 12, 471 (2001).
130. Y. I. Kim, L. Fluckiger, M. Hoffmann, L. Lartaud-Idjouadiene, J. Attkinson, and P. Maincent, *Br. J. Pharmacol.* 120, 399 (1997).
131. H. Rafati, A. G. A. Coombes, J. Adler, J. Holland, and S. S. Davis, *J. Controlled Rel.* 43, 89 (1997).
132. M. Ueda and J. Kreuter, *J. Microencapsul.* 14, 593 (1997).
133. M. Ueda, A. Iwata, and J. Kreuter, *J. Microencapsul.* 15, 361 (1998).
134. H. Murakami, M. Kobayaschi, H. Takeuchi, and Y. Kawashima, *J. Controlled Rel.* 67, 29 (2000).
135. H. M. Redhead, S. S. Davis, and L. Illum, *J. Controlled. Rel.* 70, 353 (2001).
136. H. Chen and R. Bodmeier, *J. Controlled Rel.* 12, 223 (1990).
137. F. Koosha, R. H. Müller, and C. Washington, *J. Pharm. Pharmacol.* 39, 136P (1987).
138. F. Koosha, R. H. Müller, S. S. Davis, and M. C. Davies, *J. Controlled Rel.* 12, 223 (1990).
139. L. A. Guzman, V. Labhasetwar, C. Song, Y. Jang, A. M. Lincoff, R. Levy, and E. J. Topol, *Circulation* 94, 1441 (1996).
140. C. X. Song, V. Labhasetwar, H. Murphy, X. Qu, W. R. Humphrey, R. J. Shebuski, and R. J. Levy, *J. Controlled Rel.* 43, 197 (1997).
141. C. Song, V. Labhasetwar, X. Cui, W. R. T. Underwood, and R. J. Levy, *J. Controlled Rel.* 54, 201 (1998).
142. V. Labhasetwar, C. Song, W. R. Humphrey, R. J. Shebuski, and R. J. Levy, *J. Pharm. Sci.* 87, 1229 (1998).
143. M. M. Gaspar, D. Blanco, M. E. M. Cruz, and M. J. Alonso, *J. Controlled Rel.* 52, 53 (1998).
144. R. Arshady, *J. Controlled Rel.* 17, 1 (1991).
145. H. Fessi, F. Puisieux, J. Ph. Devissaguet, N. Ammoury, and S. Benita, *Int. J. Pharm.* 55, R1 (1989).
146. N. Ammoury, H. Fessi, J. P. Devissaguet, F. Puisieux, and S. Benita, *J. Pharm. Sci.* 79, 763 (1990).
147. S. E. Dunn, A. G. A. Coombes, M. C. Garnett, S. S. Davis, M. C. Davies, and L. Illum, *J. Controlled Rel.* 44, 65 (1997).
148. J.-C. Leroux, E. Alléman, E. Doelker, and R. Gurny, *Eur. J. Pharm. Biopharm.* 41, 14 (1995).
149. P. Calvo, J. L. Vila-Jato, and M. J. Alonso, *J. Pharm. Sci.* 85, 530 (1996).
150. P. Calvo, A. Sánchez, J. Martinez, M. I. López, M. Calonge, J. C. Pastor, and M. J. Alonso, *Pharm. Res.* 13, 311 (1996).
151. T. Grovender, S. Stolnik, M. C. Garnett, L. Illum, and S. S. Davis, *J. Controlled Rel.* 57, 171 (1999).
152. S. Stolnik, M. C. Davies, L. Illum, S. S. Davis, M. Buostta, and M. Vert, *J. Controlled Rel.* 30, 57 (1994).
153. H. Ibrahim, C. Bindschaedler, E. Doelker, P. Buri, and R. Gurny, *Int. J. Pharm.* 87, 239 (1992).

154. E. Allémann, R. Gurny, and E. Doelker, *Int. J. Pharm.* 87, 247 (1002).
155. E. Allémann, E. Doelker, and R. Gurny, *Eur. J. Pharm. Biopharm.* 39, 13 (1993).
156. J. C. Leroux, R. Cozens, J. L. Roesel, B. Galli, F. Kubel, E. Doelker, and R. Gurny, *J. Pharm. Sci.* 84, 1387 (1995).
157. J. C. Leroux, E. Allémann, F. De Jaeghere, E. Doelker, and R. Gurny, *J. Controlled Rel.* 39, 339 (1996).
158. K. Sugibayashi, Y. Moromoto, T. Nadai, and Y. Kato, *Chem. Pharm. Bull.* 25, 3433 (1977).
159. K. Sugibayashi, M. Akimoto, Y. Moromoto, T. Nadai, and Y. Kato, *Chem. Pharm. Bull.* 26, 3087 (1980).
160. J. M. Gallo, C. T. Hung, and D. G. Perrier, *Int. J. Pharm.* 22, 63 (1984).
161. P. T. Gupta, C. T. Hung, and D. G. Perrier, *Int. J. Pharm.* 33, 137 (1986).
162. P. T. Gupta, C. T. Hung, and D. G. Perrier, *Int. J. Pharm.* 33, 147 (1986).
163. P. T. Gupta, C. T. Hung, F. C. Lam, and D. G. Perrier, *Int. J. Pharm.* 43, 167 (1988).
164. P. T. Gupta, F. C. Lam, and C. T. Hung, *Int. J. Pharm.* 51, 253 (1989).
165. P. T. Gupta and C. T. Hung, *Int. J. Pharm.* 59, 57 (1990).
166. P. T. Gupta, C. T. Hung, and F. C. Lam, *J. Microencapsul.* 6, 147 (1989).
167. J. J. Burger, E. Tomlinson, E. M. A. Mulder, and J. G. McVie, *Int. J. Pharm.* 23, 333 (1985).
168. M. I. Papisov, V. Y. Salvelyev, V. B. Sergienko, and V. P. Torchilin, *Int. J. Pharm.* 40, 201 (1987).
169. M. I. Papisov and V. P. Torchilin, *Int. J. Pharm.* 40, 207 (1987).
170. H. F. M. Cremers, J. Feien, G. Kwon, Y. H. Bae, S. W. Kim, H. P. M. Noteborn, and J. G. McVie, *J. Controlled Rel.* 11, 167 (1990).
171. V. Schäfer, H. v. Briesen, H. Rübsamen-Waigmann, A. M. Steffan, C. Royer, and J. Kreuter, *J. Microencapsul.* 11, 261 (1994).
172. T. Yoshioka, M. Hashida, S. Muranishi, and H. Sezaki, *Int. J. Pharm.* 8, 131 (1981).
173. Y. Tabata and Y. Ikada, *Pharm. Res.* 6, 422 (1989).
174. E. E. Hassan, R. C. Parish, and J. M. Gallo, *Pharm. Res.* 9, 390 (1992).
175. Y. Morimoto and S. Fujimoto, *CRC Crit. Rev. Therap. Drug Carrier Syst.* 2, 19 (1968).
176. P. T. Gupta and C. T. Hung, *J. Microencapsul.* 6, 427 (1989).
177. R. Arshady, *J. Controlled Rel.* 14, 111 (1990).
178. J. J. Marty, R. C. Oppenheim, and P. P. Speiser, *Pharm. Acta Helv.* 53, 17 (1978).
179. R. C. Oppenheim, *Int. J. Pharm.* 8, 217 (1981).
180. H. G. Bungenberg de Jong and H. R. Kruyt, *Proc. Koninkl. Akad. Wetensch.* 32, 849 (1929).
181. R. C. Oppenheim, J. J. Marty, and N. F. Steward, *Austr. J. Pharm. Sci.* 7, 113 (1978).
182. C. C. Boag, R. C. Oppenheim, P. Montague, and G. R. Birchall, *Asian J. Pharm. Sci.* 2, 27 (1980).
183. C. C. Oppenheim, N. F. Steward, L. Gordon, and H. M. Patel, *Drug Devel. Ind. Pharm.* 8, 531 (1982).
184. M. S. El-Samaligy and P. Rohdewald, *J. Pharm. Pharmacol.* 35, 537 (1983).
185. H.-J. Krause and P. Rohdewald, *Pharm. Res.* 2, 239 (1985).
186. G. Mukherji, R. S. R. Murthy, and B. D. Miglani, *Int. J. Pharm.* 65, 1 (1990).
187. V. Schäfer, H. von Briesen, R. Andreesen, A.-M. Steffan, C. Royer, S. Tröster, J. Kreuter, and H. Rübsamen-Waigmann, *Pharm. Res.* 9, 541 (1992).
188. A. Bender, V. Schäfer, A. M. Steffan, C. Royer, J. Kreuter, H. Rübsamen-Waigmann, and H. v. Briesen, *Res. Virol.* 145, 215 (1994).
189. A. K. Zimmer, P. Maincent, P. Thouvenot, and J. Kreuter, *Int. J. Pharm.* 110, 211 (1994).
190. A. K. Zimmer, H. Zerbe, and J. Kreuter, *J. Controlled Rel.* 32, 57 (1994).
191. A. K. Zimmer, P. Chetoni, M. F. Saettone, H. Zerbe, and J. Kreuter, *J. Controlled Rel.* 33, 31 (1995).
192. C. Coester, K. Langer, H. von Briesen, and J. Kreuter, *J. Microencapsul.* 17, 187 (2000).
193. K. Langer, C. Coester, C. Weber, H. von Briesen, and J. Kreuter, *Eur. J. Pharm. Biopharm.* 49, 303 (2000).
194. C. Weber, C. Coester, J. Kreuter, and K. Langer, *Int. J. Pharm.* 194, 91 (2000).
195. C. Coester, J. Kreuter, H. von Briesen, and K. Langer, *Int. J. Pharm.* 196, 147 (2000).
196. C. Weber, J. Kreuter, and K. Langer, *Int. J. Pharm.* 196, 197 (2000).
197. C. Weber, S. Reiss, and K. Langer, *Int. J. Pharm.* 211, 67 (2000).
198. J. Kreuter, *Pharm. Acta Helv.* 58, 217 (1983).
199. S. M. Moghimi, A. C. Hunter, and J. C. Murray, *Pharmacol. Rev.* 53, 283 (2001).
200. S. Stolnik, S. E. Dunn, M. C. Garnett, M. C. Davies, A. G. A. Coombes, D. C. Taylor, M. P. Irving, S. C. Purkiss, T. F. Tardos, S. S. Davis, and L. Illum, *Pharm. Res.* 11, 1800 (1994).
201. D. Bazile, C. Prud'Homme, M.-T. Bassoullet, M. Marlard, G. Spenlehauer, and M. Veillard, *J. Pharm. Sci.* 84, 493 (1995).
202. M. T. Peracchia, R. Gref, Y. Minamitake, A. Domb, N. Lotan, and R. Langer, *J. Controlled Rel.* 46, 223 (1977).
203. I. L. Shin, S. Y. Kim, Y. M. Lee, C. S. Cho, and Y. K. Sung, *J. Controlled Rel.* 51, 1 (1998).
204. S. Y. Kim, I. L. Shin, Y. M. Lee, C. S. Cho, and Y. K. Sung, *J. Controlled Rel.* 51, 13 (1998).
205. T. Görner, R. Gref, D. Michenot, F. Sommer, M. N. Tran, and E. Dellacherie, *J. Controlled Rel.* 57, 259 (1999).
206. F. De Jaeghere, E. Allémann, J. Feijen, T. Kissel, E. Doelker, and R. Gurny, *J. Drug Target.* 8, 143 (2000).
207. T. Jung, A. Breitenbach, and T. Kissel, *J. Controlled Rel.* 67, 157 (2000).
208. M. T. Peracchia, C. Vauthier, F. Puisieux, and P. Couveur, *J. Biomed. Mater. Res.* 34, 317 (1997).
209. M. T. Peracchia, E. Fattal, D. Desmaele, M. Besnard, J. P. Noel, J. M. Gomis, M. Appel, J. d'Angelo, and P. Couveur, *J. Controlled Res.* 60, 121 (1999).
210. M. T. Peracchia, S. Harnisch, H. Pinto-Alphandary, A. Gulik, J. C. Dedieu, D. Desmaele, J. d'Angelo, R. H. Müller, and P. Couveur, *Biomaterials* 20, 1269 (1999).
211. I. Brigger, P. Chaminade, D. Desmaele, M. T. Peracchia, J. D'Angelo, R. Gurny, M. Renoir, and P. Couvreur, *Pharm. Res.* 17, 1124 (2000).
212. I. Brigger, P. Chaminade, V. Marsaud, M. Appel, M. Besnard, R. Gurny, M. Renoir, and P. Couvreur, *Int. J. Pharm.* 214, 37 (2001).
213. W. Lin, M. C. Garnett, M. C. Davies, F. Bignott, P. Ferruti, S. S. Davis, and L. Illum, *Biomaterials* 18, 559 (1997).
214. W. Lin, M. C. Garnett, E. Schacht, S. S. Davis, and L. Illum, *Int. J. Pharm.* 189, 161 (1999).
215. W. Lin, M. C. Garnett, S. S. Davis, E. Schacht, P. Ferruti, and L. Illum, *J. Controlled Rel.* 71, 117 (2001).
216. A. Rembaum, S. Margel, and J. Levy, *J. Immunol. Meth.* 24, 239 (1978).
217. S. Margel, S. Zisblatt, and A. Rembaum, *J. Immunol. Meth.* 28, 341 (1979).
218. S. Margel, *J. Polymer Sci., Polymer Chem. Ed.* 22, 3521 (1984).
219. R. Arshady, *J. Microencapsul.* 6, 1 (1989).
220. R. Arshady, *J. Microencapsul.* 6, 13 (1989).
221. A. K. Fahlvik, E. Holtz, P. Leander, U. Schroder, and J. Klaveness, *Invest. Radiol.* 25, 113 (1990).

222. P. Edmann, B. Ekman, and I. Sjöholm, *J. Pharm. Sci.* 73, 1507 (1980).
223. P. Arthursson, P. Edmann, and I. Sjöholm, *J. Pharmacol. Exp. Terap.* 234, 255 (1984).
224. P. Arthursson, I.-L. Martensson, and I. Sjöholm, *J. Pharm. Sci.* 75, 697 (1986).
225. T. Laakso, P. Stjärnkvist, and I. Sjöholm, *J. Pharm. Sci.* 76, 134 (1987).
226. T. Laakso and I. Sjöholm, *J. Pharm. Sci.* 76, 935 (1987).
227. M. Rajaonarivomy, C. Vauthier, G. Couarraze, F. Puisieux, and P. Couvreur, *J. Pharm. Sci.* 82, 912 (1993).
228. I. Aynie, C. Vauthier, H. Chacun, E. Fattal, and P. Couvreur, *Antisense Nucl. Acid Drug Devel.* 9, 301 (1999).
229. P. Calvo, C. Remunan-Lopez, J. L. Villa-Jato, and M. J. Alonso, *Pharm. Res.* 14, 1431 (1997).
230. K. A. Janes, M. P. Fresneau, A. Marazula, A. Fabra, and M. J. Alonso, *J. Controlled Rel.* 73, 255 (2001).
231. H. Q. Mao, K. Roy, V. L. Troung-Le, K. A. Janes, K. Y. Lin, Y. Wang, J. T. August, and K. W. Leong, *J. Controlled Rel.* 70, 399 (2001).
232. A. F. Thünemann and S. General, *J. Controlled Rel.* 75, 237 (2001).
233. K. Akiyoshi, S. Deguchi, N. Moriguchi, S. Yamaguchi, and J. Sunamoto, *Macromolecules* 26, 3062 (1993).
234. K. Akiyoshi, I. Taniguchi, H. Fukui, and J. Sunamoto, *Eur. J. Pharm. Biopharm.* 42, 286 (1996).
235. T. Eldem, P. Speiser, and A. Hincal, *Pharm. Res.* 8, 47 (1991).
236. M. R. Gasco, S. Morel, and R. Carpignano, *Eur. J. Pharm. Biopharm.* 38, 7 (1992).
237. R. Cavalli, O. Carputo, and M. R. Gasco, *Int. J. Pharm.* 89, R9 (1993).
238. R. Cavalli, E. Mareng, O. Carputo, E. Ugazio, and M. R. Gasco, *J. Dispersion Sci. Technol.* 17, 717 (1996).
239. S. Morel, E. Ugazio, R. Cavalli, and M. R. Gasco, *Int. J. Pharm.* 132, 259 (1996).
240. S. Morel, E. Terreno, E. Ugazio, S. Aime, and M. R. Gasco, *Eur. J. Pharm. Biopharm.* 45, 157 (1998).
241. B. Siekmann and K. Westensen, *Pharm. Pharmacol. Lett.* 3, 194 (1994).
242. B. Siekmann and K. Westensen, *Pharm. Pharmacol. Lett.* 3, 225 (1994).
243. R. H. Müller, H. Weyhers, A. zur Mühlen, A. Dingler, and W. Mehnert, *Pharm. Ind.* 59, 423 (1997).
244. W. Mehnert, A. zur Mühlen, A. Dingler, H. Weyhers, and R. H. Müller, *Pharm. Ind.* 59, 511 (1997).
245. R. H. Müller, A. Dingler, H. Weyhers, A. zur Mühlen, and W. Mehnert, *Pharm. Ind.* 59, 614 (1997).
246. R. H. Müller, K. Mäder, and S. Gohla, *Eur. J. Pharm. Biopharm.* 50, 161 (2000).
247. S. Brandriss, G. Borchard, J. Kreuter, and S. Margel, *Reactive Polym.* 25, 111 (1995).
248. G. Borchard, S. Brandriss, J. Kreuter, and S. Margel, *J. Drug Target.* 2, 61 (1994).
249. R. H. Müller, "Colloidal Carriers for Controlled Drug Delivery and Targeting." Wissenschaftl. Verlagsges. GmbH, Stuttgart, 1990.
250. K. S. Soppimath, T. M. Aminabhavi, A. P. Kulkarni, and W. A. Rudzinski, *J. Controlled Rel.* 70, 1 (2001).
251. C. Washington, *Int. J. Pharm.* 58, 1 (1990).
252. S. D. Tröster and J. Kreuter, *Int. J. Pharm.* 45, 91 (1988).
253. J. Kreuter, in "Topics in Pharmaceutical Sciences 1983" (D. D. Breimer and P. Speiser, Eds.), p. 359. Elsevier, Amsterdam/New York/Oxford, 1983.
254. J. Vora, N. Bapat, and M. Boroujerdi, *Drug Devel. Ind. Pharm.* 19, 759 (1993).
255. J. Martin, P. Macchi, M. Hoffmann, and P. Maincent, *J. Pharm. Belg.* 49, 498 (1994).
256. E. Benoit, O. Prot, P. Maincent, and J. Bessière, *Pharm. Res.* 11, 585 (1994).
257. T. Harmia, J. Kreuter, P. P. Speiser, T. Boye, R. Gurny, and A. Kubis, *Int. J. Pharm.* 33, 187 (1986).
258. B. Magenheim and S. Benita, *S.T.P. Pharma Sci.* 1, 221 (1991).
259. J. Kreuter, S. N. Mills, S. S. Davis, and C. G. Wilson, *Int. J. Pharm.* 16, 105 (1983).
260. V. H. K. Li, R. W. Wood, J. Kreuter, T. Harmia, and J. R. Robinson, *J. Microencapsul.* 3, 213 (1986).
261. J. Kreuter, *Pharm. Acta Helv.* 58, 242 (1983).
262. T. Blunk, D. F. Hochstrasser, J.-C. Sanchez, B. W. Müller, and R. H. Müller, *Electrophoresis* 14, 1357 (1993).
263. P. G. Waser, U. Müller, J. Kreuter, S. Berger, K. Munz, E. Kaiser, and B. Pfluger, *Int. J. Pharm.* 39, 213 (1987).
264. E. M. Gipps, R. Arshady, J. Kreuter, P. Groscurth, and P. P. Speiser, *J. Pharm. Sci.* 75, 256 (1986).
265. E. M. Gipps, P. Groscurth, J. Kreuter, and P. P. Speiser, *J. Pharm. Sci.* 77, 208 (1988).
266. V. Lenaerts, P. Couvreur, D. Christiaens-Leyh, E. Joiris, M. Roland, B. Rollmann, and P. Speiser, *Biomater.* 5, 65 (1984).
267. M. Stein and E. Hamacher, *Int. J. Pharm.* 80, R11 (1992).
268. P. Couvreur and C. Vauthier, *J. Controlled Rel.* 17, 187 (1990).
269. R. H. Müller, C. Lherm, J. Herbot, and P. Couvreur, *Biomater.* 11, 590 (1990).
270. R. H. Müller, C. Lherm, J. Herbot, T. Blunk, and P. Couvreur, *Int. J. Pharm.* 84, 1 (1992).
271. M. Vert, S. Li, and H. Garreau, *J. Controlled Rel.* 16, 5 (1991).
272. T. G. Park, *J. Controlled Rel.* 30, 161 (1994).
273. M. Vert, G. Schwach, R. Engel, and J. Coudane, *J. Controlled Rel.* 53, 85 (1998).
274. J. Kreuter, M. Nefzger, E. Liehl, R. Czok, and R. Voges, *J. Pharm. Sci.* 72, 1146 (1983).
275. J. Kreuter, R. Mauler, H. Gruschkau, and P. P. Speiser, *Exp. Cell Biol.* 44, 12 (1976).
276. J. Kreuter and E. Liehl, *J. Pharm. Sci.* 70, 367 (1981).
277. J. Kreuter, *Adv. Drug Deliv. Rev.* 7, 71 (1991).
278. M. Nefzger, J. Kreuter, R. Voges, E. Liehl, and R. Czok, *J. Pharm. Sci.* 73, 1309 (1984).
279. J. Kreuter, U. Müller, and K. Munz, *Int. J. Pharm.* 55, 39 (1989).
280. L. Araujo, M. Sheppard, R. Löbenberg, and J. Kreuter, *Int. J. Pharm.* 176, 209 (1999).
281. P. Jani, G. W. Halbert, J. Langridge, and A. T. Florence, *J. Pharm. Pharmacol.* 41, 809 (1989).
282. L. Illum and S. S. Davis, in "Drug Targeting" (P. Burri and A. Gumma, Eds.), p. 65. Elsevier, Amsterdam, 1985.
283. C. J. van Oss, *Annu. Rev. Microbiol.* 32, 19 (1978).
284. H. Maeda and Y. Matsumura, *CRC Crit. Rev. Ther. Drug Carrier Syst.* 6, 193 (1989).
285. L. W. Seymour, *CRC Crit. Rev. Ther. Drug Carrier Syst.* 9, 135 (1992).
286. R. Duncan, S. Dimitrijev, and E. G. Evagorou, *S.T.P. Pharma Sci.* 6, 237 (1996).
287. D. J. Wilkins and P. A. Myers, *Brit. J. Exp. Pathol.* 47, 568 (1966).
288. K. J. Widder, A. E. Senyei, and D. F. Ranney, *Adv. Pharmacol. Chemother.* 16, 213 (1979).
289. A. Ibrahim, P. Couvreur, M. Roland, and P. Speiser, *J. Pharm. Pharmacol.* 35, 59 (1983).
290. J. L. Arias, V. Gallardo, S. A. Gómez-Lopera. R. C. Plaza, and A. V. Delago, *J. Controlled Rel.* 77, 309 (2001).
291. K. J. Widder, A. E. Senyei, H. Ovida, and P. Y. Paterson, *J. Pharm. Sci.* 70, 387 (1981).
292. L. Illum, P. D. E. Jones, J. Kreuter, R. W. Baldwin, and S. S. Davis, *Int. J. Pharm.* 17, 65 (1983).
293. A. Roland, D. Bourel, B. Genetet, and R. Le Verge, *Int. J. Pharm.* 39, 179 (1987).
294. L. Illum, P. D. E. Jones, R. W. Baldwin, and S. S. Davis, *J. Pharmacol. Exp. Ther.* 230, 733 (1984).
295. L. Illum and S. S. Davis, *J. Pharm. Sci.* 72, 1086 (1993).

296. D. Leu, B. Manthey, J. Kreuter, P. P. DeLuca, and P. Speiser, *J. Pharm. Sci.* 73, 1433 (1984).
297. L. Illum, I. M. Hunneyball, and S. S. Davis, *Int. J. Pharm.* 34, 145 (1986).
298. L. Illum, S. S. Davis, R. H. Müller, E. Mak, and P. West, *Life Sci.* 40, 367 (1987).
299. S. D. Tröster, U. Müller, and J. Kreuter, *Int. J. Pharm.* 61, 85 (1990).
300. S. D. Tröster, K. H. Wallis, R. H. Müller, and J. Kreuter, *J. Controlled Rel.* 20, 247 (1992).
301. G. Borchard and J. Kreuter, *J. Drug Target.* 1, 15 (1993).
302. L. Araujo, R. Löbenberg, and J. Kreuter, *J. Drug Target.* 5, 373 (1999).
303. L. Illum and S. S. Davis, *Life Sci.* 40, 1553 (1987).
304. A. E. Gulyaev, S. E. Gelperina, I. N. Skidan, A. S. Antropov, G. Ya. Kivman, and J. Kreuter, *Pharm. Res.* 16, 1564 (1999).
305. A. Friese, E. Seiller, G. Quack, B. Lorenz, and J. Kreuter, *Eur. J. Pharm. Biopharm.* 49, 103 (2000).
306. R. Reszka, P. Beck, I. Fichtner, M. Hentschel, J. Richter, and J. Kreuter, *J. Pharmacol. Exp. Therap.* 280, 232 (1997).
307. J. Lode, I. Fichtner, J. Kreuter, A. Berndt, J. E. Diederichs, and R. Reszka, *Pharm. Res.* 18, 1613 (2001).
308. G. Borchard and J. Kreuter, *Pharm. Res.* 13, 1055 (1996).
309. M. Lück, W. Schröder, S. Harnisch, K. Thode, T. Blunk, B.-R. Paulke, M. Kresse, and R. H. Müller, *Electrophoresis* 18, 2961 (1997).
310. S. Harnisch and R. H. Müller, *Electrophoresis* 19, 349 (1998).
311. A. Gessner, C. Olbrich, W. Schröder, O. Kayser, and R. H. Müller, *Int. J. Pharm.* 214, 87 (2001).
312. J. Kreuter, D. Shamenkov, V. Petrov, P. Ramge, K. Cychutek, C. Koch-Brandt, and R. Alyautdin, *J. Drug. Target.* 10, 317 (2002).
313. J. Kreuter, U. Täuber, and V. Illi, *J. Pharm. Sci.* 68, 1443 (1979).
314. J. H. Eldridge, C. J. Hammond, J. A. Meulbroek, J. K. Staas, R. M. Gilley, and T. R. Tice, *J. Controlled Rel.* 11, 205 (1990).
315. J. U. Jani, G. W. Halbert, J. Langridge, and A. T. Florence, *J. Pharm. Pharmacol.* 42, 821 (1990).
316. J. U. Jani, A. T. Florence, and D. E. McCarcy, *Int. J. Pharm.* 84, 245 (1992).
317. J. U. Jani, D. E. McCarthy, and A. T. Florence, *Int. J. Pharm.* 86, 239 (1992).
318. A. T. Florence, A. M. Hillary, N. Hussein, and J. U. Jani, *J. Drug Target.* 3, 65 (1995).
319. A. M. Hillary and A. T. Florence, *Int. J. Pharm.* 132, 123 (1996).
320. P. Maincent, P. Thouvenot, C. Amicabile, M. Hoffmann, J. Kreuter, P. Couvreur, and J. P. Devissaguet, *Pharm. Res.* 9, 1534 (1992).
321. J. Kreuter and H. R. Hartmann, *Oncology* 40, 363 (1983).
322. P. Couvreur, L. Roblot-Treupel, F. Puisieux, E. Fattal, A. Andremont, N. Chiannilkulchai, and F. Brasseur, in "Topics in Pharmaceutical Sciences 1989" (D. D. Breimer, D. J. A. Crommelin, and K. K. Mitra, Eds.), p. 651. Elsevier, Amsterdam, 1989.
323. N. Chiannilkulchai, Z. Driouich, J. P. Benoit, A. L. Parodi, and P. Couvreur, *Selective Cancer Therap.* 5, 1 (1989).
324. A. C. de Verdiere, C. Dubernet, F. Nemati, E. Soma, M. Appel, J. Ferte, S. Bernard, F. Puisieux, and P. Couvreur, *Brit. J. Cancer* 76, 198 (1997).
325. C. E. Soma, C. Dubernet, G. Barrat, F. Nemati, M. Appel, J. S. Benita, and P. Couvreur, *Pharm. Res.* 16, 1710 (1999).
326. C. E. Soma, C. Dubernet, D. Bentonila, J. S. Benita, and P. Couvreur, *Biomaterials* 21, 1 (2000).
327. C. E. Soma, C. Dubernet, G. Barrat, J. S. Benita, and P. Couvreur, *J. Controlled Rel.* 68, 283 (2000).
328. R. Fernadez-Urrusuno, E. Fattal, D. Porquet, J. Feger, and P. Couvreur, *Toxicol. Appl. Pharmacol.* 130, 272 (1995).
329. R. Fernadez-Urrusuno, E. Fattal, J. M. Rodrigeues, Jr., J. Feger, P. Bedossa, and P. Couvreur, *J. Biomed Mater. Res.* 31, 401 (1996).
330. S. Gibaud, J. P. Andreux, C. Weingarten, M. Renaud, and P. Couvreur, *Eur. J. Cancer* 30A, 820 (1994).
331. S. Gibaud, M. Demoy, J. P. Andreux, C. Weingarten, B. Gouritin, and P. Couvreur, *J. Pharm. Sci.* 83, 944 (1996).
332. M. Demoy, S. Gibaud, J. P. Andreux, C. Weingarten, B. Gouritin, and P. Couvreur, *Pharm. Res.* 14, 463 (1997).
333. C. Verdun, P. Couvreur, H. Vranckx, V. Lenaerts, and M. Roland, *J. Controlled Rel.* 3, 205 (1986).
334. J. Kattan, J. P. Droz, P. Couvreur, J.-P. Marino, A. Boutan-Laroze, P. Rougier, P. Brault, H. Vranckx, J.-M. Grognet, X. Morge, and H. Sancho-Garnier, *Invest. New Drugs* 10, 191 (1992).
335. S. Mitra, U. Gaur, P. C. Gosh, and A. N. Maitra, *J. Controlled Rel.* 74, 317 (2001).
336. H. S. Yoo, J. E. Oh, K. H. Lee, and T. G. Park, *Pharm. Res.* 16, 1114 (1999).
337. H. S. Yoo, K. H. Lee, J. E. Oh, and T. G. Park, *J. Controlled Rel.* 68, 419 (2000).
338. K. Avgoustakis, A. Beletsi, Z. Panagi, P. Klepetsanis, A. G. Karydas, and D. S. Ithakissios, *J. Controlled Rel.* 79, 123 (2002).
339. S. Feng and G. Huang, *J. Controlled Rel.* 71, 53 (2001).
340. D.-B. Chen, T. Yang, W.-L. Lu, and Q. Zhang, *Chem. Pharm. Bull.* 49, 1444 (2001).
341. B. Stella, S. Arpicco, M. T. Peracchia, D. Dismaele, J. Hoebeke, M. Renoir, J. D'Angelo, L. Cattel, and P. Couvreur, *J. Pharm. Sci.* 89, 1452 (2000).
342. H. Pinto-Alphandary, A. Andremont, and P. Couvreur, *Int. J. Antimicrob. Agents* 13, 155 (2000).
343. R. Gaspar, V. Préat, and M. Roland, *Int. J. Pharm.* 68, 11 (1991).
344. T. K. M. Mbela, J. H. Poupaert, and P. Dumont, *Int. J. Pharm.* 79, 29 (1992).
345. M. Fouarge, M. Dewulf, P. Couvreur, M. Roland, and H. Vranckx, *J. Microencapsul.* 6, 29 (1989).
346. C. Losa, P. Calvo, E. Castro, J. L. Villa-Jato, and M. J. Alonso, *J. Pharm. Pharmacol.* 43, 548 (1991).
347. M.-E. Page-Clisson, H. Pinto-Alphandary, M. Ourevitch, A. Andremont, and P. Couvreur, *J. Controlled Rel.* 56, 23 (1998).
348. V. D. Labhasetwar and A. K. Dorle, *J. Controlled Rel.* 12, 113 (1990).
349. V. Schäfer, J. Kreuter, H. Rübsamen-Waigmann, S. Gerte, and H. v. Briesen, *Clin. Diagnost. Virol.* 1, 279 (1994).
350. R. Löbenberg, L. Araujo, and J. Kreuter, *Eur. J. Pharm. Biopharm.* 44, 127 (1997).
351. R. Löbenberg, J. Maas, and J. Kreuter, *J. Drug Target.* 5, 171 (1997).
352. J. Kreuter, *Adv. Drug Deliv. Rev.* 47, 65 (2001).
353. U. Schroeder, H. Schroeder, and B. A. Sabel, *Life Sci.* 66, 495 (2000).
354. U. Schroeder, P. Sommerfeld, S. Ulrich, and B. A. Sabel, *J. Pharm. Sci.* 87, 1305 (1998).
355. R. N. Alyautdin, E. B. Tezikov, P. Ramge, D. A. Kharkevich, D. J. Begley, and J. Kreuter, *J. Microencapsul.* 15, 67 (1998).
356. J. Kreuter, V. E. Petrov, D. A. Kharkevich, D. A., and R. N. Alyautdin, *J. Controlled Rel.* 49, 81 (1997).
357. R. N. Alyautdin, A. Reichel, R. Löbenberg, P. Ramge, J. Kreuter, and D. J. Begley, *J. Drug Target.* 9, 209 (2001).
358. C. Michel, M. Aprahamian, L. Defontaine, P. Couvreur, and C. Damgé, *J. Pharm. Pharmacol.* 43, 1 (1991).
359. C. Damgé, C. Michel, M. Aprahamian, P. Couvreur, and J. P. Devissaguet, *J. Controlled Rel.* 13, 233 (1990).
360. M. Roques, C. Damgé, C. Michel, C. Staedel, G. Crémel, and P. Hubert, *Diabetes* 41, 451 (1992).
361. C. Damgé, H. Vranckx, P. Balschmidt, and P. Couvreur, *J. Pharm. Sci.* 86, 1403 (1977).
362. C. Damgé, M. Aprahamian, G. Balboni, A. Hoeltzel, V. Andrieu, and J. P. Devissaguet, *Int. J. Pharm.* 36, 121 (1987).
363. H. O. Alpar, W. N. Field, K. Hayes, and D. A. Lewis, *J. Pharm. Pharmacol.* 41 (Suppl), 50P (1990).

364. J. C. Leroux, R. Cozens, J. L. Roesel, B. Galli, E. Doelker, and R. Gurny, *Pharm. Res.* 13, 485 (1996).
365. S. Sakuma, N. Suzuki, H. Kikuchi, K. Hiwatari, K. Arikawa, A. Kishida, and M. Akashi, *Int. J. Pharm.* 149, 93 (1997).
366. E. Mathiowitz, J. S. Jacob, Y. S. Jong, G. P. Carino, D. E. Chickering, P. Chaturvedi, C. A. Santos, K. Vijayaraghavan, S. Montgomery, M. Bassett, and C. Morelli, *Nature* 386, 410 (1997).
367. K. Roy, H. Q. Mao, S.-K. Huang, and K. E. Leong, *Nature Med.* 5, 387 (1999).
368. N. Ammouri, H. Fessi, J.-P. Devissaguet, M. Allix, M. Plotkine, and R. G. Bouline, *J. Pharm. Pharmacol.* 42, 558 (1990).
369. N. Ammouri, H. Fessi, J.-P. Devissaguet, M. Dubrasquet, and S. Benita, *Pharm. Res.* 8, 101 (1991).
370. N. Hussain, P. U. Jani, and A. T. Florence, *Pharm. Res.* 14, 613 (1997).
371. R. W. Wood, V. H.-K. Li, J. Kreuter, and J. R. Robinson, *Int. J. Pharm.* 23, 175 (1985).
372. R. Gurny, *Pharm. Acta Helv.* 56, 130 (1981).
373. R. Gurny, T. Boye, and H. Ibrahim, *J. Controlled Rel.* 2, 353 (1985).
374. H. Ibrahim, C. Bindschaedler, E. Doelker, P. Buri, and R. Gurny, *Int. J. Pharm.* 77, 211 (1991).
375. R. Diepold, J. Kreuter, P. Guggenbühl, and J. R. Robinson, *Int. J. Pharm.* 54, 149 (1989).
376. L. Marchal-Heussler, D. Sirbat, M. Hoffmann, and P. Maincent, *Pharm. Res.* 10, 386 (1993).
377. C. Losa, L. Marchal-Heussler, F. Orallo, J. L. Villa Jato, and M. J. Alonso, *Pharm. Res.* 10, 80 (1993).
378. A. Zimmer and J. Kreuter, *Adv. Drug Delivery Rev.* 16, 61 (1995).
379. R. J. Mrsny, *J. Drug Target.* 7, 1 (1999).
380. G. Lambert, E. Fattal, and P. Couvreur, *Adv. Drug Deliv. Rev.* 47, 99 (2001).
381. G. Godard, A. S. Boutorine, E. Saison-Behmoaras, and C. Helene, *Eur. J. Biochem.* 232, 404 (1995).
382. H. P. Zobel, J. Kreuter, D. Werner, C. R. Noe, G. Kümel, and A. Zimmer, *Antisense Nucl. Acid Drug Devel.* 7, 483 (1997).
383. S. Hirosue, B. G. Müller, R. C. Mulligan, and R. Langer, *J. Controlled Rel.* 79, 231 (2001).
384. C. Kneuer, M. Sameti, E. G. Haltner, T. Schiestel, H. Schirra, H. Schmidt, and C. M. Lehr, *Int. J. Pharm.* 196, 257 (2000).
385. H.-P. Zobel, D. Werner, M. Gilbert, C. R. Noe, F. Stieneker, J. Kreuter, and A. Zimmer, *J. Microencapsul.* 16, 501 (1999).
386. G. Schwab, I. Duoux, C. Chavany, and C. Helene, *Ann. Oncol.* 5 (Suppl. 4), 55 (1994).
387. M. Junghans, J. Kreuter, and A. Zimmer, *Biochim. Biophys. Acta* 1544, 177 (2001).
388. M. Berton, P. Turelli, D. Trono, C. A. Stein, E. Alléman, and R. Gurny, *Pharm. Res.* 18, 1096 (2001).
389. M. D. Blanco and M. J. Alonso, *Eur. J. Pharm. Biopharm.* 43, 287 (1997).
390. J. C. Gautier, J. L. Grangier, A. Barbier, P. Dupont, D. Dussossoy, G. Pastor, and P. Couvreur, *J. Controlled Rel.* 20, 67 (1992).
391. L. Polato, M. Benedetti, L. Callegaro, and P. Couvreur, *J. Drug Target.* 2, 53 (1994).
392. S. Gibaud, C. Rousseau, C. Weingarten, P. Favier, L. Douay, J. P. Andreux, and P. Couvreur, *J. Controlled Rel.* 52, 131 (1998).
393. J. Kreuter, *Vaccine Res.* 1, 93 (1992).
394. J. Kreuter, in "New Generation Vaccines: The Role of Basic Immunology" (G. Gregoriadis, B. McCormack, A. C. Allison, and G. Poste, Eds.), p. 73. Plenum Press, New York, 1993.
395. F. Stieneker and J. Kreuter, in "Formulation and Delivery of Proteins and Peptides" (J. Cleland and R. Langer, Eds.), ACS Symposium Series, No. 567, p. 306. American Chemical Society, Washington, DC, 1994.
396. J. Kreuter, in "Vaccine Design: The Subunit and Adjuvant Approach" (M. F. Powell and M. J. Newman, Eds.), p. 463. Plenum, New York, 1995.
397. J Kreuter, in "Vaccine Adjuvants" (D. T. O'Hagan, Ed.), p. 105. Humana, Totowa, NJ, 2000.
398. J. Kreuter and I. Haenzel, *Infect. Immunity* 19, 667 (1978).
399. J. Kreuter and E. Liehl, *Med. Microbiol. Immunol.* 165, 111 (1978).
400. J. Kreuter, E. Liehl, U. E. Berg, M. Soliva, and P. P. Speiser, *Vaccine* 6, 256 (1988).
401. F. Stieneker, J. Löwer, and J. Kreuter, *Vaccine Res.* 2, 111 (1993).
402. F. Stieneker, G. Kersten, L. van Bloois, D. J. A. Crommelin, S. L. Hem, J. Löwer, and J. Kreuter, *Vaccine* 13, 45 (1995).
403. J. Kreuter and P. P. Speiser, *J. Pharm. Sci.* 6, 1624 (1976).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanoparticles for Live-Cell Dynamics

Xiao-Hong Nancy Xu, Rudrax N. Patel

Old Dominion University, Norfolk, Virginia, USA

CONTENTS

1. INTRODUCTION

Noble metal nanoparticles have been widely used as nanoprobes to determine the receptors on the cell surface using transmission electron microscopy (TEM) since the 1970s [1, 2]. Because cell death occurs under the vacuum of TEM, no kinetic information on the interactions between the nanoparticles with live cells was learned. As detection means have advanced, nanoparticles have been used for ultrasensitive sensing of biomolecules (DNA, proteins) in buffer solutions [3–7], and, most recently, nanoparticles have been used as nanoprobes to measure the real-time size transformation of membrane pores in live cells [8, 9]. Nanoparticles prepared with semiconductors and polymers have also been widely explored for the study of biomolecules in buffer solution [10, 11] and in live cells [12–15].

Recent developments in nanoscience and nanotechnology demonstrate the possibility of transforming biological and material research. For instance, the nanometer-scale, size-dependent, optical properties of noble metal nanoparticles have the potential to lead to new techniques for real-time live-cell imaging with sub-100-nm spatial resolution and millisecond to femtosecond time resolution. These unique capabilities could revolutionize our fundamental understanding of such important phenomena as membrane transport, protein dynamics, protein–protein interactions, and intracellular and intercellular signaling in live cells. Numerous approaches have been reported for exploring the inherent promise in the application of nanoparticle optical principles to biological problems. This review presents an overview of the research on the dynamics of biological processes in live cells or live organisms (e.g., microtubules) using nanoparticles. Three major nanoparticles, including noble metal nanoparticles (Au, Ag), semiconductor quantum dots (QDs), and fluorophores incorporated with polymer matrix nanoparticles (FluoSphere beads, PEBBLES), have been used for the study of the dynamics of biological processes in live cells and will be the focus of this review. We offer a brief summary of past and current research on this topic and provide a future outlook of this subject.

2. SYNTHESIS AND CHARACTERIZATION OF NANOPARTICLES

2.1. Noble Metal Nanoparticles

The synthesis of Au and Ag nanoparticles [16, 17] and silver-enhanced gold nanoparticles (SEGNPs) [8, 18] has been well described. Typically, the sizes of Au and Ag nanoparticles were governed by controlling the ratio of $HAuCl_4$ or $AgNO_3$ to reducing agents (trisodium citrate or tannic acid). Various sizes of SEGNPs were prepared using small colloidal gold-nucleating cores (6 nm) and a commercially available silver enhancement kit [8, 18]. These nanoparticles have their unique surface plasmon resonance characteristics and show a size dependence of absorption bands and surface plasmon resonance spectra [18–21]. Thus, these nanoparticles have been widely characterized using TEM, ultraviolet–visible (UV–vis) spectroscopy, and dark-field microscopy and spectroscopy (Fig. 1). These nanoparticles have been offered as nonbleaching multicolor optical probes for real-time tracking of cellular events for an unlimited time frame (hours/days).

The optical properties (colors) of Au and Ag nanoparticles are attributable to the surface plasmon resonance of the nanoparticles and depend on the size and shape of the nanoparticles and the dielectric constant of the embedding medium of the nanoparticles as described by Mie theory with quasistatic regime [19–21]. Unlike the bulk plasmon, the surface plasmon of nanoparticles is directly excited by propagating light waves (electromagnetic waves). This allows the nanoparticles to absorb and scatter specific wavelengths of light selectively. Therefore, one could tune the color of the nanoparticles by carefully selecting the nanoparticle

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (181–192)

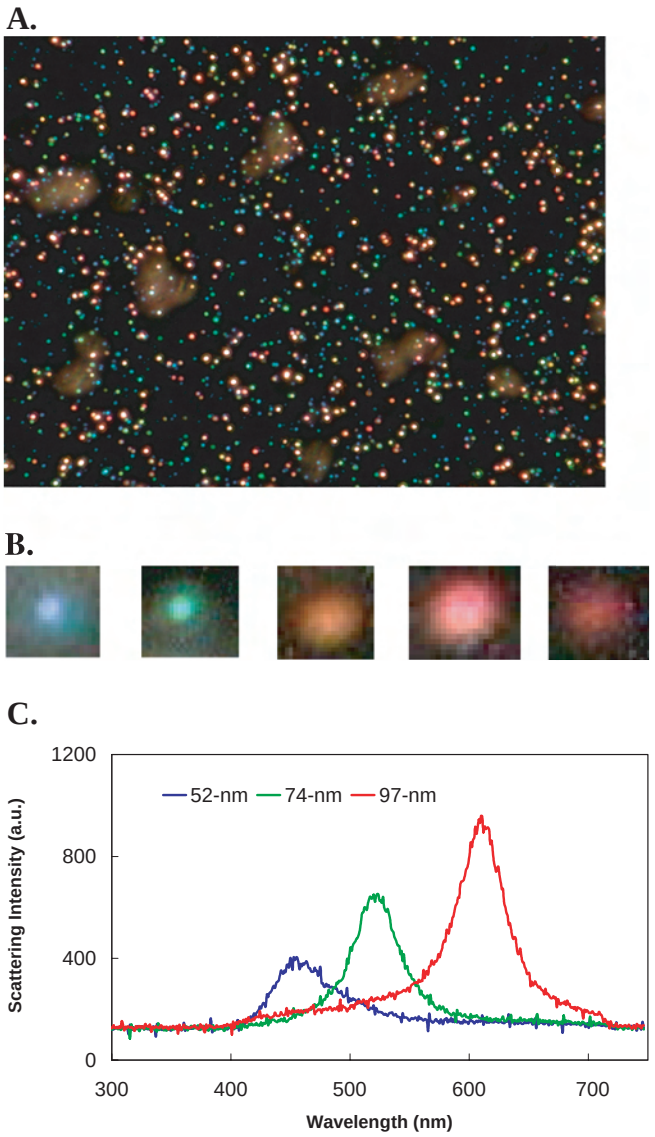


Figure 1. Representative full frame of real-color images of SEGNPs (A) and close-up real-color images (B) and surface plasmon resonance spectra (C) of single SEGNPs with sizes of 52 nm (blue color), 74 nm (green color), and 97 nm (red color) [8]. The sizes of the nanoparticles were characterized using TEM (JEOL). The images (A–B) and spectra (C) were recorded using a dark-field microscope equipped with a color digital camera (Coolpix 990) and LN-back-illuminated CCD camera coupled with SpectraPro-150 (B), respectively. These data demonstrate that it is possible to simultaneously size the nanoparticles using their colors while following the membrane transport in real time. The optical images of nanoparticles look larger than their actual sizes because of the optical diffraction limit (~200 nm). Reprinted with permission from [8], X.-H. Xu et al., *Nano Lett.* 2, 175 (2002). © 2002, American Chemical Society.

sizes and maintaining the nanoparticle shape (sphere) and embedding medium at a constant. These unique properties allowed qualitative analysis of these multicolor nanoparticles and demonstrated the possibility of distinguishing in real time the different sizes of nanoparticles in solution using dark-field microscopy and spectroscopy (Fig. 1). These nanoparticles have served as multicolor nanoprobes for measuring the real-time size transformation of membrane transporters and membrane permeability (Figs. 2–5). The optical properties of nanoparticles are preserved after functionalization of the surface of Au and Ag nanoparticles using biological molecules (antibodies) [22, 23]. These labeled nanoparticles have been used to follow the motion of single components on live cells [24–26] and to measure the real-time dynamics of ligand–receptor interactions on live cells (Fig. 5) [22, 23]. Taken together, these unique properties show the possibility of using nanoparticles for the study and characterization of molecular channels, the following of membrane pump machinery and assembly, and the interaction of several functional protein molecules on/in living cells simultaneously.

Au nanoparticles attached with fluorophore molecules (fluorescein derivatives) have been developed for the study of the flux rate of nitric oxide (NO) in live cells [14]. The orientation of fluorophore molecules on the surface of Au nanoparticles changed in the presence of NO in live cells. This led to a change in fluorescence intensity of the fluorophore molecules. This sensing scheme is based on the fluorescence spectroscopic properties of fluorophore molecules on Au nanoparticles rather than the optical scattering properties of the nanoparticles themselves. Thus, this type of nanoparticle is unable to resist photobleaching.

2.2. Semiconductor Quantum Dots

Semiconductor (CdSe) QDs 1–10 nm in diameter have been synthesized [27, 28]. The methods for the synthesis of the QDs have been continuously improved [29]. Water-soluble QDs are now commercially available and have been prepared by capping CdSe QDs with ZnS and silica-coating the CdSe/ZnS [10, 11, 30–32]. These QDs offer narrow, tunable, symmetric fluorescence emission spectra and are relatively photochemically stable. The emission wavelengths of the QDs show size dependence [27–34]. The sizes of these QDs are comparable with the sizes of proteins. In addition, the fluorescence emission of the QDs allows them to be visualized inside live cells more easily and clearly than that of the surface plasmon resonance of Au and Ag nanoparticles. Nevertheless, QDs are not as photochemically stable as noble metal nanoparticles. It is evident that the emission intensity of the QDs decreases substantially after labeling with biological molecules (e.g., antibodies) [11]. This affects the sensitivity of the QDs for bioanalysis. Furthermore, QDs suffer an on–off blinking phenomenon that may an obstacle for their use as effective probes for real-time dynamics measurements in live cells. Moreover, the toxicity of the QDs may become a major obstacle to be studied and overcome before they become popular nanoprobes for the study of live-cell dynamics and disease diagnosis in humans.

2.3. Polymer-Based Fluorophore Nanoparticles

Nanoparticles prepared by the incorporation of fluorophores with a polymeric matrix have been developed to circumvent the drawbacks of bare fluorophore molecules for the study of the dynamics of subcellular events in live cells [14, 15].

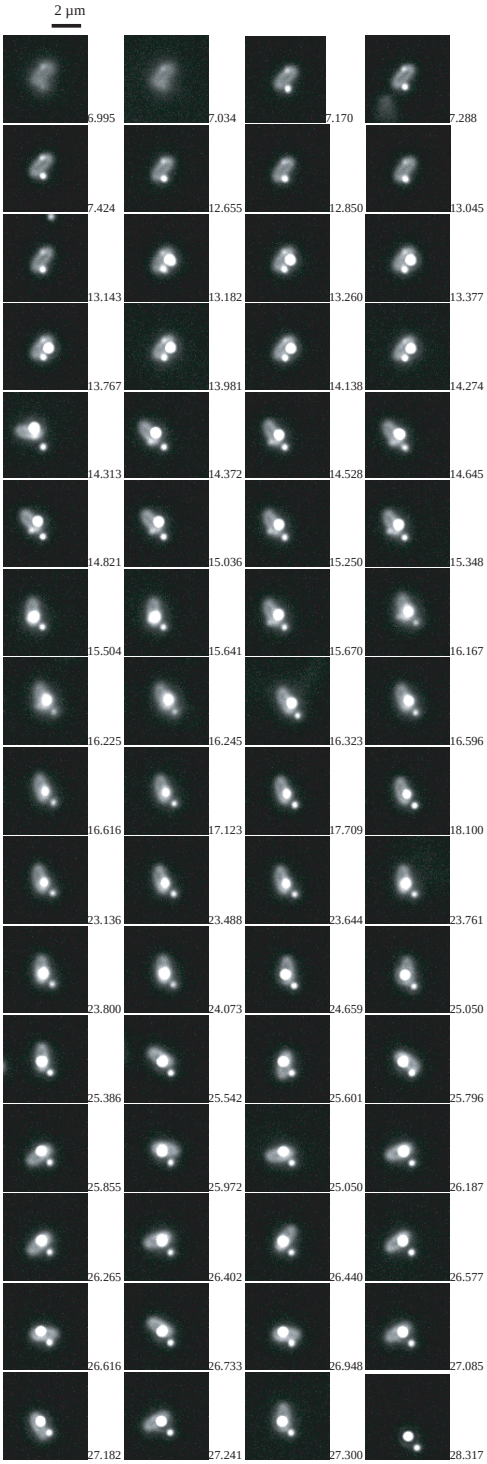


Figure 2. Representative snapshots of ΔABM cell uptake and efflux of SEGNPs using dark-field microscopy through a coverslip with a 100 × objective, equipped with a PID 1030 × 1300-pixel CCD camera (Micromax, 5 MHz-Interline), at 100 ms exposure time with a temporal resolution of 1.17 s [8]. The time shown in each image represents real time with 1 min as a unit. A blue nanoparticle (∼50 nm in diameter) participated in several uptake/efflux events in the ΔABM cell: first uptake at 7.170 min with efflux at 14.313 min; second uptake at 25.050 min and efflux at 25.386 min; third uptake at 25.601 min and efflux at 25.855 min; fourth uptake at 26.948 min and efflux at 27.085 min; and fifth uptake at 27.182 min and efflux at 27.241 min. The bright yellow particle was taken up at 13.182 min and extruded at 28.317 min by the same ΔABM cell. This yellow particle (∼80 nm) is larger than the blue nanoparticle (∼50 nm) and therefore appears brighter. The larger size made it harder for the nanoparticle to enter the cell. The cellular membrane absorbs and scatters the light. This leads to a reduction in the illumination intensity inside the cell. Thus, the nanoparticles look significantly dimmer inside the cells than outside the cells. This allows us to distinguish if the nanoparticles are inside or outside the cells. The nanoparticles appear to stick out of the cellular membrane. This is because the size of the nanoparticles and the thickness of the cell membrane are under the optical diffraction limit and the nanoparticles scatter light much more efficiently and appear brighter than the cell membrane. The gray scale was set at 262–500 electron counts. The color of the nanoparticles was directly observed and confirmed with spectroscopy. Samples were prepared by directly mixing 20 μL of the cell solution (10 × dilution of absorbance at 600 nm = 0.1) with 20 μL of a 2.60-pM SEGNPs solution. The timer was started at the point of mixing. Reprinted with permission from [8], X.-H. Xu et al., *Nano Lett.* 2, 175 (2002). © 2002, American Chemical Society.

Two primary types of fluorophore polymer-based nanoparticles are fluorescence FluoSphere beads (20–100 nm) [35] and PEBBLEs (probes encapsulated by biologically localized embedding) (20–200 nm) [14, 15]. Like fluorophores, these polymer-based fluorophore nanoparticles use the fluorescence properties of fluorophores as sensing elements and hence are unable to overcome the inherent photobleaching of fluorophores even though the inner reference (free fluorophores) has been used to deduct the possible photobleaching of the fluorophores. This affects the sensitivity, stability, and lifetime of these nanoparticles. Thus, in general, these nanoparticles have a shorter lifetime than that of QDs and Au and Ag nanoparticles.

FluoSphere beads are prepared by loading polystyrene nanospheres with a variety of fluorescent dyes. These fluorescence nanospheres provide emission spectra ranging from the near ultraviolet to the near infrared and are able to be used in multicomplex analysis [35]. These nanospheres are also easily functionalized for sensing the specific target molecules. This property circumvents the limitation of reactive dyes that easily disrupt the function of biomolecules upon conjugation. Furthermore, these fluorescence nanospheres offer highly intensive fluorescence emission and can be visualized at the single-nanoparticle level. These nanospheres also provide a longer shelf life and can be monitored for a longer lifetime than free fluorophore molecules in solution. Moreover, the polymeric matrix protects the embedded fluorophores from nonspecific interactions and reduces the toxicity of the fluorophores to live cells.

PEBBLEs have been prepared by incorporation of fluorescence indicators with a polyacrylamide matrix and have been used to measure ion flux rates in live cells as described by Kopelman and co-workers [14, 15]. Unlike fluorescent indicators, the fluorophore molecules in PEBBLEs are protected by the polymeric matrix in order to avoid the interaction of fluorophore molecules with proteins in live cells and to eliminate the possibility of false results and the toxicity of dyes. This approach allows the dyes to stay in the live cells for a relatively longer time and to be used for monitoring the dynamics of ion fluxes in live cells more effectively.

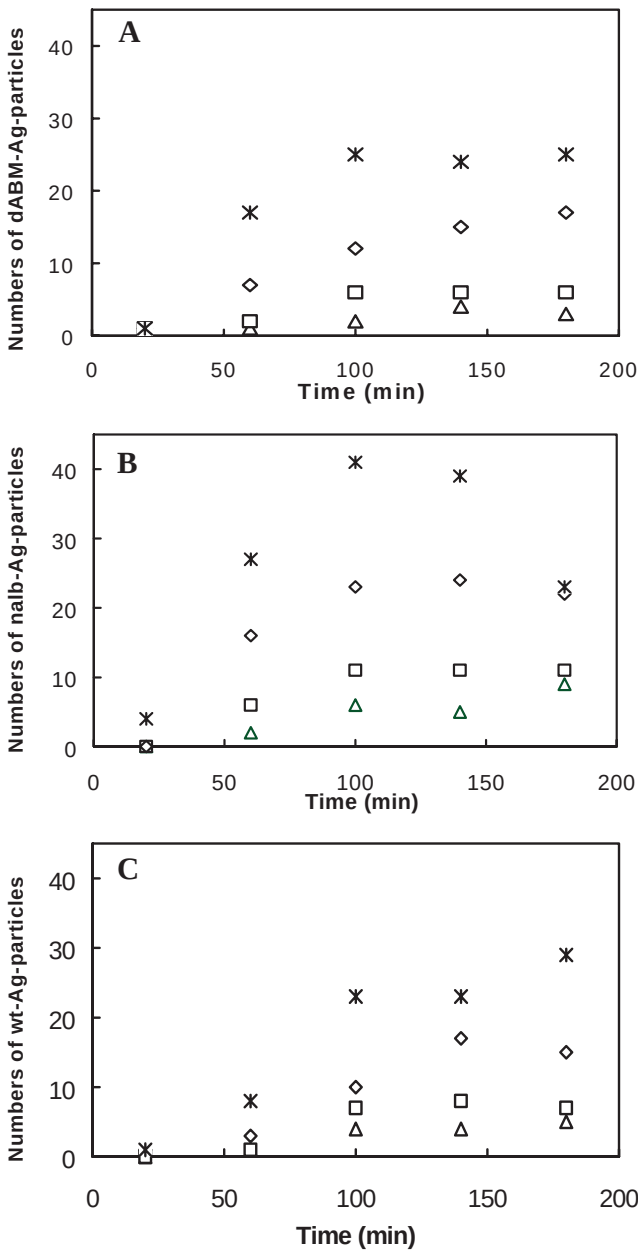


Figure 3. Representative plots of the number of cells with SEGNPs versus incubation time: (A) ΔABM, (B) nalB-1, (C) WT [8]. Samples were prepared by incubating the cell solution (absorbance at 600 nm = 0.1) with the nanoparticle solutions at 0.26 (Δ), 0.52 (□), 1.30 (◇), and (*) pM in a vial while the timer was simultaneously activated. At each corresponding time, a sample of 40 μL was transferred to the microchamber and imaged directly by dark-field microscopy using a CCD camera and color digital camera (Fig. 2). Ten images taken from each sample similar to those shown in Figure 2 were analyzed by counting the number of cells with and without nanoparticles. The total number of cells analyzed for each cell type was 400. Reprinted with permission from [8], X.-H. Xu et al., *Nano Lett.* 2, 175 (2002). © 2002, American Chemical Society.

3. DYNAMICS OF LIVE CELLS

The study of the dynamics of live cellular membranes using single-particle tracking (SPT) has been described in several excellent review articles [36–39]. The use of nanoparticles as probes for the study of the real-time dynamics of subcellular events can be tracked back to the 1980s, when Nuydens and co-workers demonstrated the detection of microtubule-dependent intracellular motility using 20- to 40-nm Au nanoparticles through NANOVID (nanoparticle video ultramicroscopy) [40]. Nuydens and co-workers explored the application of Au nanoparticles for the study of molecular biology by the conjugation of Au nanoparticles with specific ligands. They demonstrated that Au nanoparticles conjugated with ligands could be used to determine the intracellular motility, cell membrane dynamics, receptor translocation, internalization, and intracellular routing [41]. These pioneering efforts realized the potential of nanoparticles as an essential tool for the study of live cells. Nuydens and co-workers provided a new and improved automatic technique for the study of intracellular mobility of live cells using the latest tracking algorithms and image processing hardware of that time [42]. They found the results collected from live-cell experiments in astonishingly close agreement with the experiments in a buffer solution containing microtubules and a kinesin-like protein extract. The development of this technique led to the successful determination of distributions of jump time, jump velocity, stop time, and orientation of the Au nanoparticles in PTK-2 cells [42]. They demonstrated that it was possible to monitor the lateral diffusion and the rearward movements of cellular protein toward the boundary between lamelloplasm and perinuclear cytoplasm in live cells using Au nanoparticles with a velocity of 0.5–1 μm/min and a diffusion coefficient of 0.1–0.2 $\mu m^2/s$. Inspired by this work, Sheetz and co-workers developed the techniques to study the motion of motor molecules and membrane proteins in live cells using nanoparticles [43–45]. By the 1990s, these techniques have been advanced to record the reorganization of individual components on live cells [45].

Since then, an array of nanoparticles, including noble metal nanoparticles and polymer-based nanoparticles incorporated with dyes, has been developed for the study of the dynamics of biochemical events in live cells. Kopelman and co-workers studied the kinetics of calcium release in live cells using nanoparticles prepared by the incorporation of the polymer matrix with fluorescent indicators (PEBBLEs) [14, 15]. They also developed Au nanoparticles attached with fluorescent dyes (fluorescein derivative) and then used these labeled nanoparticles to measure NO release in macrophages [13]. Recently, Xu and co-workers (our research group) applied the new approaches for the study of the dynamics in live cells using the unique properties of Ag and Au nanoparticles [8, 9]. We directly measured the real-time transformation of the size and dynamics of single-living-membrane transporters using nanoparticle optics and single-live-cell imaging [8]. The unique optical properties of the Ag nanoparticles allow us to monitor in real time the size transformation of the single-membrane transporters of live cells at a nanometer resolution and a millisecond temporal resolution using dark-field microscopy and spectroscopy. We also studied the ligand–receptor interaction on a single live cell in real time using SEGNPs [22].

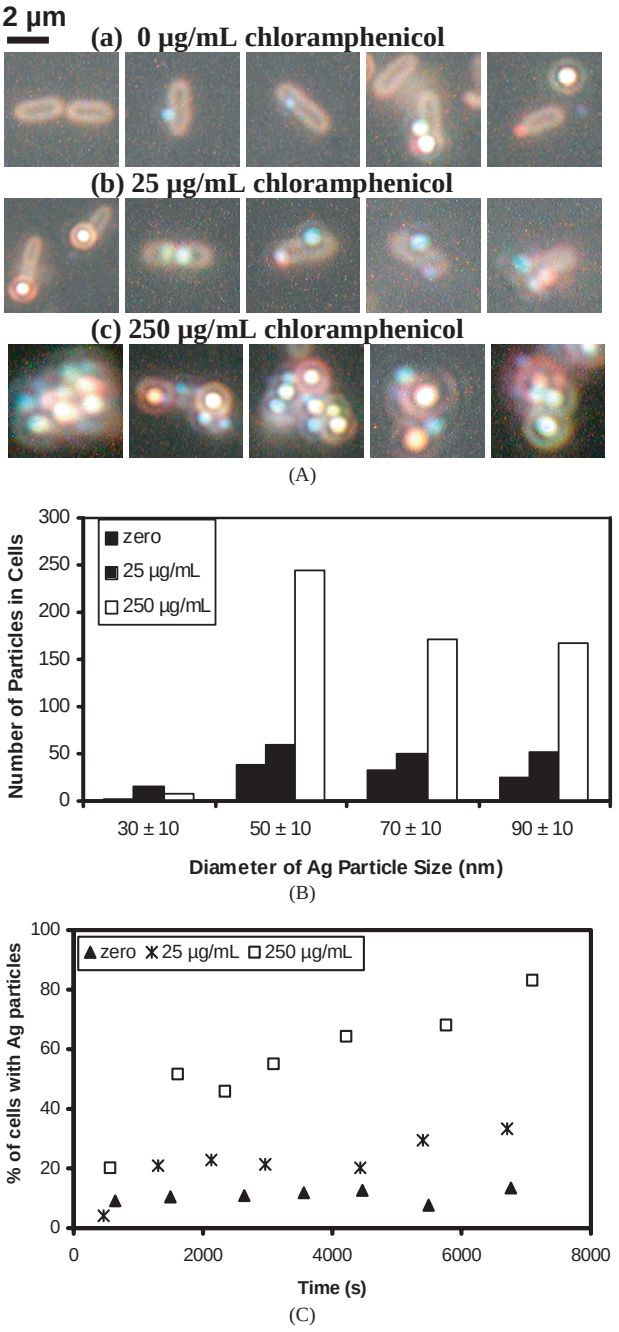


Figure 4. Direct observation of real-time size transformation of membrane pores (permeability) in WT cells [9]. (A) Representative optical images (a–c) of single cells (WT) recorded by the CCD and digital color camera through a dark-field optical microscope for 2 h. The single cell was selected from approximately 60 cells in the full-frame images. The solutions containing the WT cells ($A_{600nm} = 0.1$), 1.3 pM Ag nanoparticles, and (a) 0 μg/mL, (b) 25 μg/mL, (c) 250 μg/mL chloramphenicol, were prepared in a vial and imaged in a fresh microchannel every 15 min. (B) Representative histograms of Ag nanoparticles in WT cells versus sizes of nanoparticles acquired from 70 images (~4200 cells), recorded for 2 h, showing a great number of larger nanoparticles were presented in the membrane at the higher concentration of chloramphenicol. The solutions were the same as those in (A). (C) Representative plots of percentage of WT cells with Ag nanoparticles versus time from solutions in (A). Images for each solution were acquired at each particular time and about 600 cells were analyzed for each point.

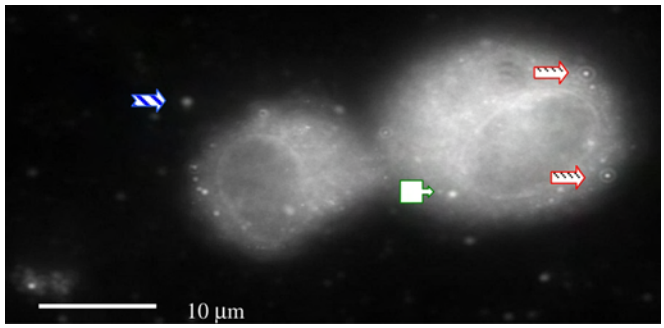


Figure 5. Real-time images of single IgG molecules binding with single ZZ-T (ZZ is a short derivative of protein A) anchored on single live fibroblast cells (L929) cultured directly on a microscope coverslip [22, 23]. The interaction of single IgG with ZZ-T on live-cell surfaces was monitored in real time. (a) Binding IgG (), (b) unbound IgG (), (c) already bound IgG ().

Nanoparticles have been developed as delivery systems and imaging probes at a brisk pace in recent years. Many researchers have described the development of nanoparticles for the study of microtubules [39, 46–48]. The relative motions of the microtubules have been measured using Au nanoparticles labeled with monoclonal tubulin antibodies [47]. These tools have been extended to study the motor-driven microtubule transport, measure the forces required for ligand–receptor interactions, and monitor the kinetics of ligand–receptor interactions on the surface of the microtubules [47, 48]. Semiconductor QDs have also been used to study ligand–receptor interactions in buffer solution and have shown a great potential for use as probes for the study of the kinetics of interactions in live cells [10, 11].

In summary, the study of the dynamics of subcellular events in live cells using nanoparticles has drawn much attention and appears to have gained momentum in recent years as research in nanoscience and nanotechnology progresses.

4. APPLICATIONS OF LIVE-CELL DYNAMICS

4.1. Intracellular Cell Motility

Nanoparticles have been used to study and characterize motion in the cytoplasm or cellular membrane of live cells. The nanoparticles have been designed and conjugated with a specific antigen or antibody to monitor the motion of specific components on the membrane of live cells [24–26]. The movement of these nanoparticles was detected and followed using a video camera as the nanoparticles moved along with the components of interest and left their discerning motility path [40, 41]. Colloidal Au nanoparticles were very popular for such studies because individual nanoparticles could be followed for an extended length of time to obtain real-time mobility using dark-field optical microscopy. Single Au nanoparticles were directly visualized by optical dark-field microscopy using a video camera [49, 50]. The mobility of Au nanoparticles has been studied in PTK-2 cells [42].

The motion of single Au nanoparticles has been determined using algorithms and image processing hardware

[42]. Typical patterns of saltatory and Brownian motion were taken into account for statistical analysis. Saltatory movement was characterized by long jumps (>1 μm) and long stop times (2–5 s), whereas Brownian motion was characterized by small elongations and no stop times [51]. These criteria allow one to continuously update the distributions of the jump time, jump velocity, jump direction, stop time, and orientation on the coordination matrix. If the next direction lay within 30° of the former direction, it was considered to be linear saltatory motion, whereas if that was not the case, the former motion was considered to be a finished saltatory motion. The jump time was used to determine whether the nanoparticles were participating in energy-demanding [adenotriphosphate (ATP) utilizing movement] saltatory movement. If the motion was not saltatory, then it was determined to be Brownian random motion. The diffusion coefficient was used to assign the movements of the nanoparticles in Brownian random motion.

Bare metal nanoparticles or metal nanoparticles labeled with fluorescent dyes have been used as probes for the study of intracellular cell motility [40–42, 49, 50, 52]. Furthermore, these nanoparticles conjugated with antibodies, receptors, and proteins have been employed to determine the motility characteristics of the corresponding components on the surface of live cells or inside the cells [24–26]. This technique provides the mobility features of individual nanoparticles and allows a large group of single nanoparticles to be monitored in real time simultaneously. Unlike classical fluorescence recovery after photobleaching (FRAP) techniques that offer the average motion of a number of components [53], nanoparticles were used to study individual cellular components. Thus, the subtle and heterogeneous spatial differentiation of the motion is unmasked.

Recently, FluoSphere beads (20 nm) labeled with a Fab fragment of monoclonal antibody (W6/32) have been developed and used for the study and characterization of the motion of MHC (major histocompatibility complex) class I molecules on the surface of HeLa cells [52]. The labeled FluoSphere beads were tracked in real time using high-sensitivity fluorescence microscopy. The spots recorded at 4-s intervals by fluorescence microscopy were connected to deduce the path of mobility (Fig. 6). Typically, a higher mobility was observed from bare metal nanoparticles and a lower mobility was observed from nanoparticles incorporated with fluorescence dyes or polymer based because bare nanoparticles have less interaction with biological entities.

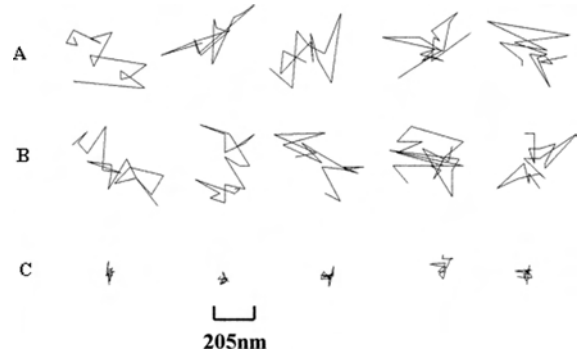


Figure 6. Tracks generated by MHC class I molecules on the surface of HeLa cells obtained using high-sensitivity fluorescence imaging [52]. The tracks of MHC class I were obtained using 20 nm FluoSphere beads labeled with PE-Fab, CML 100-Fab, and CML 30-Fab. Representative samples of the tracks obtained using PhyE-Fab (A) and CML 100-Fab (B) on cells and PhyE-Fab (C) on poly-L-lysine-coated microscope slides. Anomalous diffusion patterns were deduced from these tracks generated by linking the spots obtained at approximately 4-s intervals. Reprinted with permission from [52], P. Smith et al., *Biophys. J.* 76, 3331 (1999). © 1999, Biophysical Society.

4.2. Cell Motility

Motility and migration of cancerous cells lead to metastases and the formation of secondary tumors. Cell motility is correlated with the metastases of the cells and has been studied using phagokinetics tracks [54, 55]. Several methods have been developed to record the phagokinetic tracks of cells [55–59]. One method, based on real-time observation of phagokinetic tracks of cells, has been developed, in which single live cells are passed over a layer of markers and ingest the markers and hence individual cells leave behind a trail of blank spots that automatically register the area that the cells traveled through [55]. These trails reflect the phagokinetic tracks of cells. The advantage of this method is that it provides a rapid and automatic method for the study of cell motility that reflects both temporal and spatial information of the individual paths. The selection of markers for this method has been difficult because these markers have to be able to be uniformly coated on the surface, be ingested rapidly by cells, and endure long luminescence. Organic dyes are rapidly bleached and are unable to serve as markers. Au nanoparticles have served as the markers in this method for the study of phagokinetic tracks of cells. Due to the low-sensitivity detection means, larger Au nanoparticles (150 nm) have been used as markers for this method and hence it suffers severe drawbacks such as the lack of uniformity in both the film of Au nanoparticles and the sizes of Au nanoparticles. As ultrasensitive detection means advance, smaller Au and Ag nanoparticles (10–20 nm) will be used to circumvent these problems.

Recently, relatively photochemically stable luminescence probes, semiconductor QDs ($CdSe/ZnS/SiO_2$)(4–8 nm), have been used to study the motility of live cells [12]. These QDs are stable and soluble under physiological buffer conditions and have been used as efficient light emitters for a variety of applications [27–34]. Thin layers of QDs were deposited on a collagen-coated tissue culture substrate and then the cells were seeded on the substrate [12]. The cells and QDs were monitored in real time by using confocal fluorescence microscopy or multiphoton microscopy. As the cells migrated and ingested the QDs, the cells freed the immobilized QDs and created a trail of blank spots that clearly revealed the cell migration paths (Fig. 7). Thus, the migratory paths of tumors and normal cells were assayed simultaneously at millisecond temporal resolution. This allowed the cell motility to be calculated and the metastases of the cells to be studied. QDs have been conjugated with a wide variety of molecules for the detection of targeting molecules. Moreover, the emission spectra of QDs are size dependent [27–34]. These unique features of QDs have empowered

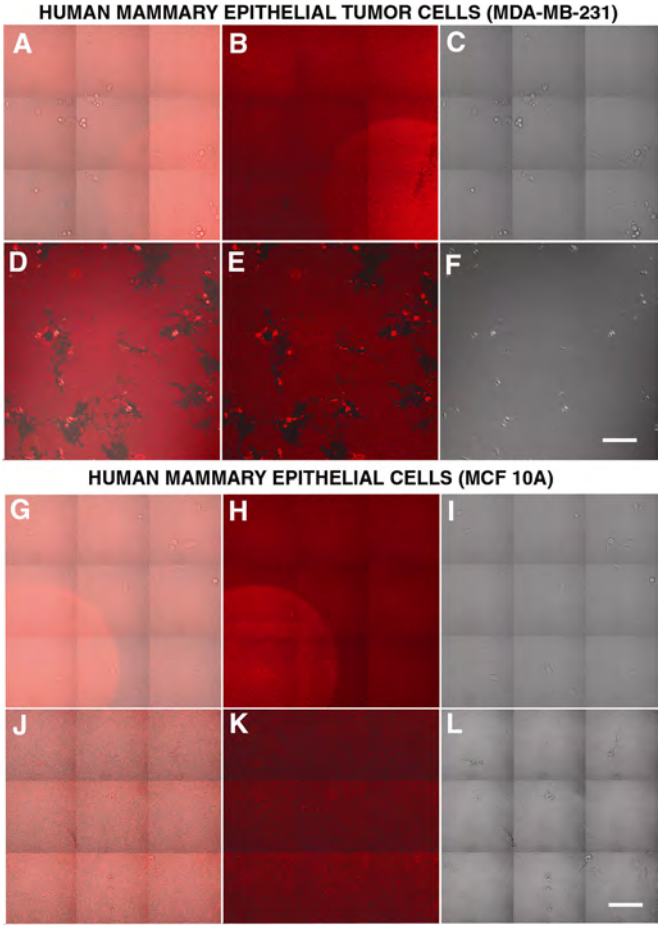


Figure 7. Migratory paths of the tumor cells determined by differential interference contrast (DIC) and fluorescence confocal microscopy [12]. Migratory paths of the tumor cells could be determined by engulfment of QDs. Human mammary epithelial tumor cells, MDA-MB-231 (A–F), and nontumor cells, MCF 10A (G–L), were grown on collagen that had been coated with a thin layer of silanized, water-soluble, red fluorescent semiconductor QDs. The merged images (A, D, G, J) show that the cells and the layer of QDs beneath the tumor cells (B) and nontumor cells (H) were fairly uniform. After 24 h, (i) the large clearings in the QD layer were observed around the tumor cells (D, E), but not around the nontumor cells (J, K). (ii) Tumor cells filled with QDs offered intensive bright fluorescence (E), whereas nontumor cells did not provide fluorescence (K). (iii) Some tumor cells were crawling into the underlying collagen and these tumors were less distinct in the DIC image (F). The images were collected using fluorescence confocal microscopy to examine QDs (B, E, H, K) and using DIC optical microscopy to visualize the cells (C, F, I, L) where the images were collected at the optical section to show the layer of QDs. The scale bar was 200 μm. The results show that the tumor cells are more motile. Reprinted with permission from [12], W. Parak et al., *Adv. Mater.* 14, 882 (2002). © 2002. Wiley–VCH.

multicomplex analysis in buffer solution [60, 61]. This offers the possibility of the measurement of cell motility and the detection of a group of specific proteins simultaneously in real time using QDs. This also makes it feasible to study cell motility in three dimensions if the different sizes of QDs in a vertical gradient are prepared in extracellular matrix media. Studies of this kind will surely advance our understanding of cell motility and metastases.

4.3. Single-Membrane Transporters

Multidrug extrusion systems have been reported for both prokaryotes and eukaryotes [62–64]. Membrane permeability and active extrusion systems appear to play a crucial role in universal cellular defense mechanisms. For instance, the efflux pump of *Pseudomonas aeruginosa* can extrude a variety of structurally and functionally diverse antibiotics [62]. It is very likely that membrane proteins are triggered by substrates to assemble membrane transporters optimizing for the extrusion of encountering substrates. Thus, the study of the real-time size and the dynamics of membrane transporters will lead to a better understanding of the multidrug resistance mechanism.

X-ray crystallographic methods have been used to determine the sizes of membrane transporters and the cellular transport mechanism. This method is limited by the difficulties of crystallization of membrane proteins and this method cannot provide the real-time dynamics of live cells [65–69]. The metal nanoparticle with the help of molecular dynamic microscopy solves this problem to a large extent [8]. The optical properties of the noble metal nanoparticles (SEGNPs) are attributed to its surface plasmon resonance and show size dependence [19–21]. Thus, we used the color index of nanoparticles as the size index to study the real-time size transformation of membrane transporters in live cells (Figs. 1–3). The sizes of the nanoparticles transported through the live-cell membranes can be measured by directly visualizing the color of individual nanoparticles, which are correlated with the size of the nanoparticles. This study has demonstrated that nanoparticle optics showed the possibility of real-time live-cell imaging with sub-100-nm spatial resolution and millisecond time resolution. Furthermore, unlike conventional fluorescent dyes or QDs, noble metal nanoparticles resist photobleaching and blinking. These unique features allow these nanoparticles to be used for continuously monitoring the size transformation of membrane transporters and efflux dynamics in live cells for an unlimited time. Nanoparticle optics is superior to TEM and scanning probe microscopy (SPM) for the study of subcellular dynamics in live cells because TEM cannot provide real-time measurements of dynamics events in live cells and SPM offers lower temporal resolution and cannot monitor nanoparticles inside live cells or many single nanoparticles simultaneously.

Three strains of *P. aeruginosa* [70–72], wild-type (PAO4290, a wild-type cell level expression of MexAB-OprM), ΔABM (TNP076, deletion of MexAB-OprM, derivative of PAO4290), and nalB-1 (TNP030#1, overexpression of all MexA, MexB, and OprM, derivative of PAO4290), have been constructed and studied. The MexAB-OprM membrane pump was used as a working model. We investigated the roles of these membrane proteins on the extrusion system by measuring the sizes and dynamics of substrates (nanoparticles) transported by mutants of overexpression or deletion of these subunits. The sizes of single substrates (nanoparticles) transported through live-cell membranes were measured by directly visualizing the color of the individual nanoparticles and hence the sizes of the extrusion pump were determined in real time at the nanometer scale (Figs. 1–3).

Based on random-walk theory [73, 74], the transport of nanoparticles through the live-cell membrane (in a buffer solution, on the cell membrane, and inside the cells) was determined [8]. Representative snapshots of real-time images of single nanoparticles transported across live-cell membranes in Figure 2 shows all three strains of *P. aeruginosa* were uptake and efflux both blue ($\sim 50 \pm 5$ nm) and green ($\sim 70 \pm 5$ nm), and, occasionally, yellow nanoparticles ($\sim 75 \pm 5$ nm). The red nanoparticles ($\sim 95 \pm 5$ nm) tended to remain outside the cells on the cell membranes and were rarely observed inside the cells (Fig. 2). These results suggest that the membrane of three mutants of *P. aeruginosa* permeated and extruded substrates up to 80 nm in diameter.

The accumulated nanoparticles by cells showed time and concentration dependence (Fig. 3). The uptake and efflux dynamics of nanoparticles by three mutants of *P. aeruginosa* differ at 100 min of incubation time. A greater number of nalB-1 cells took up and retained nanoparticles during the first 100 min in comparison with ΔABM (MexAB-OprM-deficient mutants) and WT. At 100 min, nalB-1 cells began to extrude the nanoparticles and the number of cells with nanoparticles decreased whereas ΔABM and WT have distinguished uptake and efflux characteristics. The observation of uptake and efflux of nanoparticles by ΔABM implies that nanoparticles (substrate) may trigger the assembly of membrane transporters optimized for the extrusion of encountered substrates.

This work demonstrates the possibility of real-time tracking of membrane permeability and the observation of pump machinery of single live cells at the single-extrusion-pump level and may lead to a better understanding of multidrug resistance and designing of pump inhibitors. Furthermore, alternative molecular pump designs may be explored so that pumps may one day be engineered to selectively transport specific molecules into and out of live cells for a variety of applications (e.g., intelligent drug delivery).

4.4. Membrane Permeability and Efflux Mechanism

Real-time size transformation of membrane pores (permeability) and efflux kinetics in three mutants of *P. aeruginosa*, WT, nalB-1, and ΔABM, have been directly observed in the presence of 0, 25, and 250 μg/mL chloramphenicol using nanoparticle optics and single-live-cell imaging. The membrane permeability in *P. aeruginosa* (WT, nalB-1, ΔABM) increased as the chloramphenicol concentration increased from 0 to 250 μg/mL and as the incubation time increased (Fig. 4), indicating the real-time size transformation of membrane pores (permeability). The results suggest that the mechanism of multi-antibiotic resistance may include the induction of transformation of membrane permeability and intrinsic efflux pump machinery. This work constitutes the first direct observation of the size transformation of membrane pores (permeability) and the interplay of membrane permeability and efflux kinetics in real time at single-live-cell resolution and offers the new possibility of advancing the understanding of multidrug resistance. This study demonstrates the possibility of using nanoparticles for the study of the real-time transformation of membrane permeability in eukaryotic cells (tumor) for better understanding of multidrug resistance. We are currently developing nanoparticle optics for the study of the real-time transformation of electropermeabilization [75, 76] for a better understanding of the effect of radio frequency (RF) and electric fields on membrane permeability.

4.5. Ligand–Receptor Interactions

The kinetics and binding of the ligand–receptor interaction on live cells play an important role in a variety of crucial biological events, such as immune regulation pathways. The motions of receptors on live cells using Au nanoparticles labeled with ligands have been studied previously [24–26]. We have recently developed and applied multicolor SEGNPs and real-time diffusion single-nanoparticle/single-cell microscopy for the measurement of single ligand–receptor interactions on live cells [22, 23]. We directly observed single IgG molecules labeled with single SEGNPs binding with single antigen molecules (ZZ-T) anchored on single living cells.

The ZZ is a derivative of protein A [77] that is renowned for its specific powerful binding with IgG, which serves as an invaluable molecular linker for a variety of immunoassays and affinity chromatography. This ZZ protein implies dramatic potential to build a specific molecular architecture on synthetic and natural membranes for a variety of applications (e.g., sensor, separation, nanoassemble) and can serve as a model for fundamental thermodynamics and kinetics studies at living interfaces. The ZZ-T was constructed by fusing the ZZ protein with a transmembrane domain (T) of diphtheria toxin [78–80]. Soluble protein molecules were fused with the T domain and then the T-domain fusion proteins were anchored to the membrane by triggered pH change from pH 7.4 to 5.0. The anchoring process was easily controlled and occurred rapidly.

We have anchored ZZ-T onto live cells and studied the binding of IgG (immunoglobulin G) with the ZZ-T (fusion protein) on live cells [22]. The SEGNPs were prepared and labeled with the IgG. The binding of IgG-SEGNPs with ZZ-T anchored onto live cells has been directly measured in real time as shown in Figure 5 [22]. The bound and unbound IgG-SEGNPs were distinguished using the diffusion distances of SEGNPs. The unbound probe created a large diffusion image and appeared dim because the intensity of the nanoparticles within the exposure time is distributed over each pixel of the entire diffusion area. The bound probe shrank to a single pixel of CCD resolution and appeared as a single bright spot because the bound probe on live cells would not move around and hence a single pixel collected all scattering intensity during the entire exposure time. The blank control experiments of ZZ-T-anchored cells incubated with nanoparticles or cells in the absence of ZZ-T incubated with IgG-SEGNPs showed no significant binding [22]. This study demonstrated that the nanoparticles can be used to monitor in real time the interactions of ligands with receptors on live-cell surfaces, which will lead to a better understanding of cellular events for a variety of applications (e.g., vaccine and drug design).

Other fusion proteins, T-hIL-2 and T-mIL-3, have also been constructed by fusing human interleukin 2 (hIL-2) and murine interleukin 3 (mIL-3) to the C terminus or N terminus of the T domain [78]. We are currently measuring the real-time binding kinetics and affinity of single anchored T-IL-2 (ligand) with IL-2R (receptor) labeled with nanoparticles on single live cells under RF radiation and electric fields using single-nanoparticle/single-cell microscopy. We aim to investigate whether the binding kinetics and affinity of single IL-2 (ligands) with single receptors (IL-2R) are altered by RF radiation and electric fields due to the possible changes in the dipoles of the ligands and receptors. This study will lead to a better understanding of the mechanisms of radiotherapy and electrotherapy.

4.6. Chemical Analysis

Changes in intracellular ion concentration accompany a variety of biological processes in live cells, including transport, signaling, and enzyme function. A few ions (proton and calcium) are very important in cellular functions. The ion flux has been measured by monitoring the emission intensity of the fluorophore indicators associated with intracellular ion concentrations. This study led to a better understanding of cellular events. The major drawbacks of such fluorophore indicators were that they interacted with the proteins in the cell and gave false results [81–83]. In addition, these dyes were toxic to cells and were rapidly extruded out of the cells. This limited the observation period. Moreover, these fluorescent dyes were distributed throughout the entire cell and created a huge background noise. Nanoparticles prepared by incorporating fluorescent dyes with a polymer matrix were proven to circumvent many of these problems [14, 15]. The PEBBLE sensors (20–200 nm) were prepared by incorporation of fluorescent dye molecules into the pores of the polymer matrix. The polymer matrix provided a protective coating for the dye molecules. This minimized nonspecific binding and eliminated the toxic effect of the dyes. Hence, it improved its application for real-time cellular studies. Multiple-dye PEBBLE sensors have been prepared by the integration of more than one type of dye in the polymer matrix [14, 15]. Some of these dye molecules were used as a real-time reference. Another advantage of PEBBLE sensors is that these sensors are small enough to be biocompatible and yet large enough to be retained in the cytoplasm with no leakage out of the cells for days [14, 15].

These nanoparticles enable us to study the rate of ion release in real time, which reflects specific cellular events. For example, the rate of calcium release from the mitochondria of C6 glioma cells and SY5Y neuroblastoma cells has been studied by stimulating the mitochondrial permeability transition (MPT) with varying dosages of *m*-DNB (*meta*-dinitrobenzene, a neurotoxicant). The exposure to *m*-DNB resulted in the uncontrollable release of calcium associated with the onset of MPT. The increase in intracellular calcium led to cell apoptosis. It was found that these two types of cells had different resistance and sensitivity to *m*-DNB (Fig. 8). The fluorescence intensity of the calcium-green family of dyes increased as the calcium concentration increased (Fig. 9). The change in calcium flux was

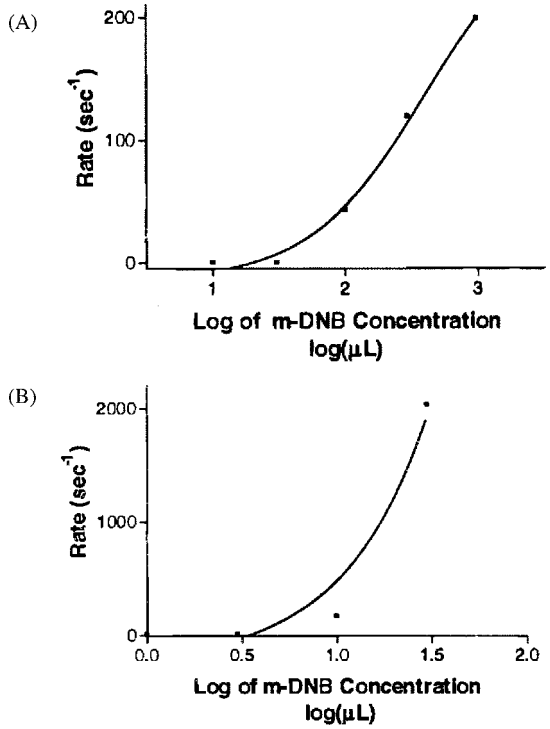


Figure 8. Direct observation of calcium release from the mitochondria of C6 glioma cells [14]. Representative plots of rate of calcium release from the mitochondria of C6 glioma cells (A) and the mitochondria of SY5Y neuroblastoma cells (B) versus the dose of *m*-DNB (*meta*-dinitrobenzene, a neurotoxicant) used to stimulate the mitochondrial permeability transition (MPT) measured by a fluorometer. Fluorescent nanoparticles (PEBBLEs) were prepared by incorporation of the dyes with the polymer matrix (polyacrylamide). Reprinted with permission from [14], H. A. Clark et al., *Anal. Chem.* 71, 4831 (1999). © 1999, American Chemical Society.

determined by measuring the changes in fluorescence intensity of the dyes embedded in the polyacrylamide matrix. These PEBBLE sensors have been used to measure the intracellular calcium fluxes in single live cells using confocal microscopy and to monitor bulk cells in real time using fluorescence spectroscopy via a fluorimeter.

Unlike PEBBLE sensors, Au nanoparticles attached with fluorophores (fluorescein derivatives) have been developed for the real-time study of NO release in macrophages [13]. The NO sensor was prepared by attaching a difluorofluorescein derivative onto Au nanoparticles [13]. This sensor measured the real-time release of NO from macrophages when macrophages were activated by LPS (lipopolysaccharide) and IFN-γ (interferon-γ). The difluorofluorescein derivatives attached perpendicularly on Au nanoparticles changed the orientation to the parallel position on the Au nanoparticle in the presence of NO. This led to a change in the fluorescence intensity. This change was used to detect the release of the NO from the live cells. Moreover, this sensor was stable toward the superoxide and hydrogen peroxide that are the byproducts of the action of LPS and IFN-γ on the cell. The advantages of these nanoparticle sensors offer a new opportunity for the study of real-time dynamics of ion flux in live cells.

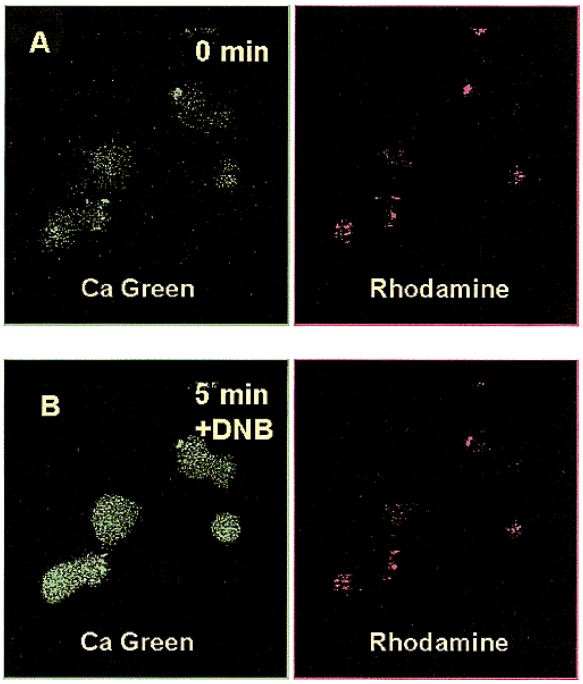


Figure 9. Direct observation of calcium release in SY5Y neuroblastoma cells. Laser scanning confocal microscopic images of *m*-DNB-stimulated calcium release in SY5Y neuroblastoma cells lipofected with calcium/sulforhodamine 101 PEBBLEs. Neuroblastoma cells were incubated in Dulbecco's phosphate buffered saline supplemented with glucose (0.5 g/L) at 37 °C and treated with 10 μM *m*-DNB. Calcium green and rhodamine fluorescence were monitored at time 0, prior to the addition of *m*-DNB (A) and 5 min following addition (B) [14]. The fluorescence intensity of the calcium green increased, which indicated that the calcium concentration level increased. As expected, the internal standard remained constant. Reprinted with permission from [14], H. A. Clark et al., *Anal. Chem.* 71, 4831 (1999). © 1999, American Chemical Society.

5. SUMMARY

Nanoparticles have shown great potential for the study of the dynamics of biological processes in live cells at the nanometer spatial resolution and millisecond time resolution. A new era in the study of the dynamics of biological events in live cells using the unique properties of nanoparticles and rapid ultrasensitive detectors (CCD cameras) have just begun. As we navigate through the wonderland of nanoscience and nanotechnology, the most noteworthy research topics surrounding the study of live-cell dynamics using nanoparticles are envisioned to be the following: (i) New tools for the nanofabrication and characterization of nanomaterials will continuously emerge and hence more new nanomaterials and new properties of nanomaterials will be discovered. Thus, more types of nanoparticles (e.g., magnetic nanoparticles, carbon-tube nanoparticles, chemically and biologically modified nanoparticles) for the study of specific biological processes (e.g., signal transduction) will be developed. (ii) Nanoparticles have been widely used as nanosensors to quantitatively and qualitatively determine species (DNA, protein, chemicals) in buffer solution [3–7, 84]. Future research will focus more on the study of biocompatible nanosensors and delivery of these nanosensors into live cells or into living organisms [85] to sense the desired biomolecules (e.g., tumor markers) for monitoring the onset and development of diseases and for understanding unclear biological mechanisms. (iii) As technologies advance, more sensitive and rapid detection means will be available. The combination of rapid and ultrasensitive detection means with unique characteristics of the nanoparticles, including nanosensing [84], multicomplexing [60, 61], nanometer resolution [9, 10], and photostability, will make real-time study of biological mechanisms at the single-molecule level feasible and hence advance our understanding of biological phenomena in real time at the nanometer scale and molecule level. We will be able to understand when and where the crucial biological events are taking place in live cells and what molecules are involved in these processes. An approach of this kind will certainly revolutionize our understanding of biology and related science and engineering. (iv) The development of nanoparticles for the study of the dynamics of biochemical events in live cells will also open up new opportunities for the study of the interactions of nanomaterials with live organisms at the molecular level. New research areas on the study of the assembly of new materials using the unique properties of live organisms for the development of self-repairing intelligent materials or the assembly of live organisms using the self-assembly of nanomaterials will appear. These research fields will revolutionize the biological sciences, materials sciences, and every aspect of science and engineering. The bright future of the exploration of live cells using nanoparticles is certain and will surely exceed our current imagination and expectation.

GLOSSARY

Charge-coupled device (CCD) A photodetector used in digital camera systems for sensitive detection of photos.

Dark-field microscopy A microscopic method that creates the dark background and the bright object with a high contrast between the object and the surrounding field [73].

Deoxyribonucleic acid (DNA) A macromolecule consists of one or two strands of linked deoxyribonucleotides.

Differential interference contrast microscopy (DIC) A microscopic method derives contrast in an unstained specimen from differences in index of refraction of specimen components [73].

Fluorescence The emission of light by a substance immediately after the absorption of energy from light of usually shorter wavelength.

Fluorescence recovery after photobleaching (FRAP) A technique to observe and quantify the movement of molecules by loading fluorescent tagged molecules, bleaching the fluorescence of a defined region of the specimen, and recovery of fluorescence in the bleached area based upon the mobility of the fluorescent molecules.

Human interleukin 2 (hIL-2) A cytokine that stimulates the immune system.

Interleukin 2 receptor (IL-2R) The protein has specific binding affinity with interleukin-2.

Metastases Pathogenic microorganisms or cancerous cells transmitted from an original site to one or more sites elsewhere in the body.

Mouse interleukin 3 (mIL-3) A peptide cytokine that stimulates haemopoiesis and activates mononuclear phagocytes and B and T lymphocytes.

Nanoparticle video ultramicroscopy (NANOVID) Tracking of the trajectories of individual microscopic particles attached to relevant molecules using optical microscopy equipped with video camera [40].

Nitric oxide (NO) A colorless gas, an important nitrogen compound, plays numerous biological roles in living organisms.

Probes encapsulated by biologically localized embedding (PEBBLES) The polymer-based fluorophore nanoparticles prepared by incorporation of fluorophores with a polymeric matrix [14, 15].

Pseudomonas aeruginosa (P. aeruginosa) A gram-negative bacterium that resists to a wide spectra of structurally and functionally unrelated substrates and is the epitome of an opportunistic pathogen of humans.

Quantum dots Inorganic nanocrystalline particles made from semiconductor materials.

Scanning probe microscopy (SPM) An experimental imaging technique maps the surface at (near) atomic resolution by scanning the surface of the sample using a sharp probe.

Silver enhanced gold nanoparticle (SEGNP) The surface of gold nanoparticles is coated with a thin layer of silver to increase the intrinsic optical dielectric constant of nanoparticle and enhance the detection sensitivity [8, 18].

Single particle tracking (SPT) A method to track the trajectories of individual microscopic particles attached to relevant molecules.

T domain: Transmembrane domain of diphtheria toxin.

Transmission electron microscope (TEM) An imaging technique takes advantage of shorter wavelengths of electrons to increase magnification and resolution.

T-ZZ A short derivative of protein A fused with amino-terminal of transmembrane domain of diphtheria toxin.

ZZ A short derivative of protein A.

ZZ-T A short derivative of protein A fused with carboxyl-terminal of transmembrane domain of diphtheria toxin.

ACKNOWLEDGMENTS

We are indebted to all the students of the Xu group for their contribution to the development of nanoparticles for the study of live-cell dynamics. We thank Taiji Nakae (Tokai University School of Medicine, Japan) for three strains of *Pseudomonas aeruginosa*. The support of this work in part by the NIH (Grant RR15057-01), DOD (Grant AFOSR-MURI), and Old Dominion University, in the form of a start-up grant, is gratefully acknowledged.

REFERENCES

1. J. E. Beesley, "Colloidal Gold: A New Perspective for Cytochemical Marking", p. 25. Oxford University Press, Oxford, UK, 1989.
2. K. Manweiler, H. Hohenberg, W. Bohn, and G. Rutter, *J. Microsc.* 126, 145 (1982).
3. C. A. Mirkin, R. L. Letsinger, R. C. Mucic, and J. J. Storhoff, *Nature* 382, 607 (1996).
4. R. Elghanian, J. J. Storhoff, R. C. Mucic, R. L. Letsinger, and C. A. Mirkin, *Science* 277, 1078 (1997).
5. T. A. Taton, G. Lu, and C. A. Mirkin, *J. Am. Chem. Soc.* 123, 5164 (2001).
6. J. J. Storhoff, R. Elghanian, R. C. Mucic, and C. A. Mirkin, *J. Am. Chem. Soc.* 120, 1959 (1998).
7. S. R. Nicewarner-Pena, R. G. Freeman, B. D. Reiss, L. He, D. J. Pena, I. D. Walton, R. Cromer, C. D. Keating, and M. J. Natan, *Science* 294, 137 (2001).
8. X.-H. Xu, J. Chen, R. Jeffers, and S. Kyriacou, *Nano Lett.* 2, 175 (2002).
9. X.-H. Xu, S. Kyriacou, W. Brownlow, Q. Wan, J. Viola, and R. Patel, submitted.
10. M. Bruchez, M. Moronne, P. Gin, S. Weiss, and A. P. Alivisatos, *Science* 281, 2013 (1998).
11. W. C. W. Chan and S. Nie, *Science* 281, 2016 (1998).
12. W. Parak, R. Boudreau, M. Le Gros, D. Gerion, D. Zanchet, C. Micheel, S. Williams, A. P. Alivisatos, and C. Larabell, *Adv. Mater.* 14, 882 (2002).
13. L. Susan, R. Barker, and R. Kopelman, *Anal. Chem.* 70, 4902 (1998).
14. H. A. Clark, M. Hoyer, M. A. Philbert, and R. Kopelman, *Anal. Chem.* 71, 4831 (1999).
15. H. Clark, R. Kopelman, R. Tjalkens, and M. Philbert, *Anal. Chem.* 71, 4837 (1999).
16. M. A. Hayat, "Colloid Gold: Principles, Methods and Applications," Vol. 1, pp. 15–27. Academic Press, New York, 1989.
17. P. C. Lee and D. Meisel, *J. Phys. Chem.* 86, 3391 (1982).
18. S. Schultz, D. Smith, J. Mock, and D. Schultz, *Proc. Natl. Acad. Sci. U.S.A.* 97, 996 (2000).
19. G. Mie, *Ann. Phys.* 25, 377 (1908).
20. C. F. Bohren and D. R. Huffman, "Absorption and Scattering of Light by Small Particles," pp. 287–380. Wiley, New York, 1983, and references therein.
21. U. Kreibig and M. Vollmer, "Optical Properties of Metal Clusters," pp. 14–123. Springer-Verlag, Berlin, 1995, and references therein.
22. R. Jeffers, D. Gillet, and X.-H. Xu, submitted.
23. R. Jeffers and X.-H. Xu, "Proceedings of the 2000 Pittsburgh Conference."
24. R. Simson, B. Yang, S. E. Moore, P. Doherty, F. S. Walsh, and K. A. Jacobson, *Biophys. J.* 74, 297 (1998).
25. C. Dietrich, B. Yang, T. Fujiwara, A. Kusumi, and K. Jacobson, *Biophys. J.* 82, 274 (2002).
26. M. De Brabander, R. Nuydens, A. Ishihara, B. Holifield, K. Jacobson, and H. Geerts, *J. Cell Biol.* 112, 111 (1991).
27. D. Gerion, F. Pinaud, S. C. Williams, W. J. Parak, D. Zanchet, S. Weiss, and A. P. Alivisatos, *J. Phys. Chem.* 105, 8861 (2001), and references therein.
28. M. Bruchez, Jr., M. Moronne, P. Gin, S. Weiss, and A. P. Alivisatos, *Science* 281, 2013–2016 (1998), and references therein.
29. Z. A. Peng and X. Peng, *J. Am. Chem. Soc.* 123, 183 (2001).
30. M. A. Hines and P. Guyot-Sionnest, *J. Phys. Chem.* 100, 468 (1996).
31. B. O. Dabbousi, J. Rodriguez-Viejo, F. V. Mikulec, J. R. Heine, H. Mattoussi, R. Ober, K. F. Jensen, and M. G. Bawendi, *J. Phys. Chem.* 101, 9463 (1997).
32. http://www.qdots.com/.
33. T. D. Lacoste, X. Michalet, F. Pinaud, D. S. Chemla, A. P. Alivisatos, and S. Weiss, *Proc. Natl. Acad. Sci. U.S.A.* 97, 9461 (2000).
34. A. P. Alivisatos, *J. Phys. Chem.* 100, 13226 (1996).
35. http://www.probes.com/servlets/product/, and references therein.
36. S. Inoué, "Video Microscopy," p. 584. Plenum, New York, 1986.
37. S. Inoué, *Methods Cell Biol.* 30, 85 (1989).
38. D. Shotton, "Electronic Light Microscopy", p. 355. Wiley, New York, 1993.

39. M. J. Saxton and K. Jacobson, *Annu. Rev. Biophys. Biomol. Struct.* 26, 373 (1997), and references therein.
40. M. De Brabander, G. Geuens, R. Nuyens, M. Moeremans, and J. De Mey, *Cytobios* 43, 273 (1985).
41. M. De Brabander, R. Nuyens, G. Geuens, M. Moeremans, and J. De Mey, *Cell Motil. Cytoskel.* 6, 105 (1986).
42. H. Geerts, M. De Brabander, R. Nuyens, G. Geuens, M. Moeremans, J. De Mey, and P. Hollenbeck, *Biophys. J.* 52, 775 (1987).
43. J. Gelles, B. J. Schnapp, and M. P. Sheetz, *Nature* 331, 450 (1998).
44. B. J. Schnapp, J. Gelles, and M. P. Sheetz, *Cell Motil. Cytoskel.* 10, 47 (1988).
45. M. P. Sheetz and S. C. Kuo, *Methods Cell Biol.* 39, 129 (1993).
46. H. Hess, J. Clemmens, J. Howard, and V. Vogel, *Nano Lett.* 2, 113 (2002).
47. Y. Wada, T. Hamasaki, and P. Satir, *Mol. Biol. Cell* 11, 161 (2002).
48. H. Hess, J. Howard, and V. Vogel, *Nano Lett.* 2, 1113 (2002).
49. R. Allen, J. Travis, N. Allen, and H. Yllmaz, *Cell Motil.* 1, 275 (1981).
50. S. Inoué, *J. Cell Biol.* 89, 346 (1981).
51. J. Freed and M. Verowitz, *J. Cell Biol.* 45, 334 (1972).
52. P. Smith, I. E. Morrison, K. Wilson, N. Fernandez, and R. Cherry, *Biophys. J.* 76, 3331 (1999).
53. G. Georgiou, S. S. Bahra, A. R. Mackie, C. A. Wolfe, P. O'Shea, S. Ladha, N. Fernandez, and R. J. Cherry, *Biophys. J.* 82, 1828 (2002).
54. A. W. Partin, J. S. Schoeniger, J. L. Mohler, and D. S. Coffey, *Proc. Natl. Acad. Sci. U.S.A.* 86, 1254 (1989).
55. G. Albrecht-Buehler, *Cell* 12, 333 (1977).
56. G. Albrecht-Buehler, *Cell* 12, 395 (1977).
57. T. T. Rajah, S. M. A. Abidi, D. J. Rambo, J. J. Dmytry, and J. T. Pento, *In Vitro Cell Dev. Biol. Anim.* 34, 626 (1998).
58. Z. Kornyei, A. Czirok, T. Vicsek, and E. Madarasz, *J. Neurosci. Res.* 61, 421 (2000), and references therein.
59. J. Yao, L. Harvath, D. L. Gilbert, and C. A. Colton, *J. Neurosci. Res.* 27, 36 (1990), and references therein.
60. M. Han, X. Gao, J. Z. Su, and S. Nie, *Nature Biotechnol.* 19, 631 (2001).
61. W. C. Chan, D. J. Maxwell, X. Gao, R. E. Bailey, M. Han, and S. Nie, *Curr. Opin. Biotechnol.* 13, 40 (2002).
62. T. Nakae, *Microbiologia* 13, 273 (1997), and references therein.
63. S. P. Cole, G. Bhardwaj, J. H. Gerlach, J. E. Mackie, C. E. Grant, K. C. Almquist, A. J. Stewart, E. U. Kurz, A. M. Duncan, and R. G. Deeley, *Science* 258, 1650 (1992).
64. M. F. Rosenberg, Q. Mao, A. Holzenburg, R. C. Ford, R. G. Deeley, and S. P. Cole, *J. Biol. Chem.* 276, 16076 (2001).
65. C. G. Tate, *FEBS Lett.* 504, 94 (2001), and references therein.
66. G. A. Scarborough, *Cell Mol. Life Sci.* 57, 871 (2000), and references therein.
67. Y. Mishima, K. Momma, W. Hashimoto, B. Mikami, and K. Murata, *Acta Crystallogr., Sect. D* 57, 884 (2001).
68. C. G. Tate, E. R. Kunji, M. Lebendiker, and S. Schuldiner, *EMBO J.* 20, 77 (2001).
69. R. L. Juliano and U. Ling, *Biochim. Biophys. Acta* 455, 152 (1976).
70. A. Ocaktan, H. Yoneyama, and T. Nakae, *J. Biol. Chem.* 272, 21964 (1997).
71. L. Guan, M. Ehrmann, H. Yoneyama, and T. Nakae, *J. Biol. Chem.* 274, 10517 (1999).
72. K. Saito, H. Yoneyama, and T. Nakae, *FEMS Microbiol. Lett.* 179, 67 (1999).
73. P. W. Atkins, "Physical Chemistry," pp. 823–905. Freeman, San Francisco, 1982.
74. A. J. Bard and L. R. Faulkner, "Electrochemical Methods Fundamentals and Applications," pp. 488–510. Wiley, New York, 1980.
75. J. Teissié, N. Eynard, B. Gabriel, and M. P. Role, *Adv. Drug Delivery Rev.* 35, 3 (1999), and references therein.
76. J. C. Weaver and Y. A. Chizmadzhev, *Bioelectrochem. Bioenerg.* 41, 35 (1996), and references therein.
77. B. Nilsson, T. Moks, B. Jansson, L. Abrahmsen, A. Elmbrad, E. Holmgren, C. Henrichson, T. A. Jones, and M. Uhlen, *Protein Eng.* 1, 107 (1987).
78. D. Liger, P. Nizard, C. Gaillard, J. C. vanderSpek, J. R. Murphy, B. Pitard, and D. Gillet, *Protein Eng.* 11, 1111 (1998).
79. P. Nizard, A. Chenal, B. Beaumelle, A. Fourcade, and D. Gillet, *Protein Eng.* 14, 439 (2001).
80. P. Nizard, D. Liger, C. Gaillard, and D. Gillet, *FEBS Lett.* 433, 83 (1998).
81. J. Slavik, "Fluorescent Probes in Cellular and Molecular Biology," CRC Press, Boca Raton, FL, 1994.
82. R. Nuccitelli, in "Methods in Cell Biology," Vol. 40, p. 368. Academic Press, San Diego, 1994.
83. W. N. Ross, *Biophys. J.* 64, 1655 (1993).
84. A. J. Haes and R. P. Van Duyne, *J. Am. Chem. Soc.* 124, 10596 (2002).
85. M. E. Akerman, W. C. Chan, P. Laakkonen, S. N. Bhatia, and E. Ruoslahti, *Proc. Natl. Acad. Sci. U.S.A.* 99, 12617 (2002).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanoparticles in Nanostructured Polymers

Lyudmila M. Bronstein

Indiana University, Bloomington, Indiana, USA

CONTENTS

1. INTRODUCTION

Nanostructured materials containing metal or semiconductor nanoparticles present an exciting area of nanoscience and nanotechnology as particle incorporation allows imparting unique properties to polymeric materials: catalytic, optical, magnetic, etc. If the polymeric matrix is already nanostructured before nanoparticle formation (i.e., contains domains of different chemical nature divided by interfaces), this reflects a further degree of nanostructural organization. Normally, the presence of such nanostructures in polymer systems (presence of interfaces) allows one to carry out a subtle control over nanoparticle growth, particle size distribution, and particle surface interactions. The above characteristics are most important in determining properties of these nanomaterials and their possible applications. Narrow particle size distribution is crucial in developing well-defined materials for any application, as optical, magnetic, and even catalytic properties strongly depend on the precise control over a particle size. This particular field of science began developing about 10 years ago when papers dealing with nanoparticle formation in microsegregated block copolymers and block copolymer micelles were first published. Nowadays this field is characterized by an avalanche of publications and fast progress in discovery of new materials. Varying the polymer systems and different approaches in nanoparticle formation, nanoparticles of various size and shape, nanotubes, nanofibers, and other nanoinclusions can be obtained in polymer nanoenvironment.

The structure of this review article is the following. All methods of nanoparticle formation in nanostructured polymers will be combined into three groups depending on the medium: nanoparticle formation (i) in solid nanostructured polymer materials, (ii) in pores or cavities of nanometer size located in polymers (this makes polymers nanostructured), (iii) and in solutions of nanostructured polymers. Each type of polymeric nanoenvironment is characterized by particular features which will be thoroughly discussed. Characterization methods most commonly used for studying nanostructured polymers containing nanoparticles will be described. At the end, the properties of nanostructured polymeric systems with nanoparticles will be briefly described.

2. SYNTHESIS

2.1. Solid Nanostructured Polymers

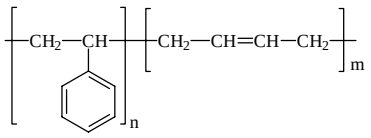
2.1.1. Microsegregated Amphiphilic Block Copolymers

An important feature of amphiphilic block copolymers is their ability to microsegregate in bulk due to incompatibility of the blocks [1–3]. Yet, for interaction with metal compounds, a chosen block (where particles will be formed) of diblock or triblock copolymer should contain functional groups. Materials of this kind can be obtained using several avenues: (i) by copolymerization of organometallic monomers with a regular monomer followed by further transformations to obtained nanoparticles in a single block, (ii) by interaction of metal complexes with functional groups of block copolymers in solutions followed by microsegregation of organometallic block copolymers in the film or in bulk, and (iii) by incorporation of organometallic compounds directly into the already preformed microdomains of the block copolymer film. Inclusion of metal complexes in di- and triblock copolymers may strongly change the properties of the corresponding blocks and influence the

ISBN: 1-58883-063-2/$35.00

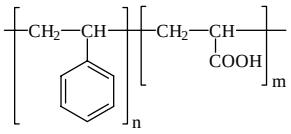
Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (193–206)

Flory–Huggins interaction parameter χ_{AB}. This influence can be especially pronounced for nonpolar flexible blocks, such as polyisoprene (PI) or polybutadiene (PB), due to the appearance of polar groups and change of chain conformation. The representative example is introduction of $Fe(CO)_3$ groups in the PB block of polystyrene-*block*-polybutadiene (PS-*b*-PB-*b*-PS, Kraton, 30 wt% PS) block copolymer due to formation of π-conjugated diolefin complexes [4].



(PS-*b*-PB)

This significantly increases adhesion of polymer film to polar surfaces (metal, glass), increases rigidity, and decreases the mobility of the PB block, though the basic morphology of this material ("cylinders in continuous phase") did not change. Similar stability of the morphology was observed also for highly asymmetric diblock ionomers where a short block (polyacrylic acid, PAA) forms spherical microdomains inside the matrix of the major component (PS) [5].



(PS-*b*-PAA)

Neither incorporation of Cd^{2+} ions nor formation of CdS nanoparticles changes the block copolymer morphology. On the contrary, the morphology of triblock copolymers strongly changed when metal complexes are introduced in the middle block [6]. If noticeable influence of metal incorporation on block copolymer morphology is rather rare, the opposite influence (i.e., the influence of nanostructures and interfaces between block copolymer domains or free surfaces of thin films) plays an important role in nanoparticle formation within such domains. Growth of Ag [7], Pd [8], and other nanoparticles inside the microsegregated layers is limited by the thickness of the layer (Fig. 1). In so doing, penetration of metal nuclei or growing particles in the adjacent nonpolar layers is negligible. If a similar particle synthesis is carried out in the thick homopolymer films, uncontrolled particle formation with broad particle size distribution is observed [7]. On the other hand, if the thickness of the film of metal-containing homopolymer diminished to the block copolymer lamellae size (about 100 nm), the free surface of the thin film behaves similarly to the interface between block copolymer domains [7].

2.1.2. Polyelectrolyte Multilayers

Such nanostructured systems are obtained by deposition of polyelectrolytes (PE) on oppositely charged supports followed by further deposition of oppositely charged (toward first layer) polyelectrolytes with a formation of multilayer polyelectrolyte systems [9]. These systems strongly resemble thin homopolymer films and lamellar block copolymers, as

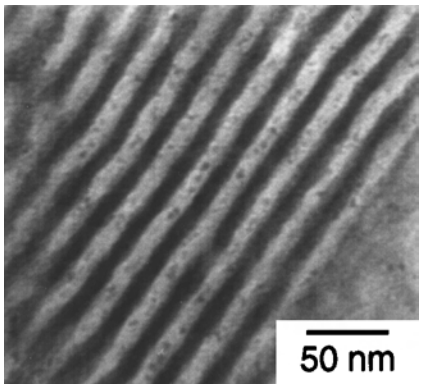
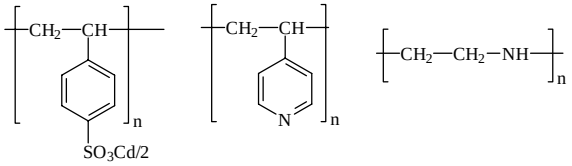


Figure 1. TEM image of the PI-*b*-P2VP/$Pd(acac)_2$ system obtained after the reduction and further stained by OsO_4 vapor. The dark and bright phases are PI lamellae stained by OsO_4 and unstained P2VP lamellae, and dark dots are the Pd nanoparticles in the P2VP phase. Reprinted with permission from [7], J. F. Ciebien et al., *New J. Chem.* 22, 685 (1998). © 1998, American Chemical Society.

the interfaces between polyelectrolyte layers limit the migration of ions to the adjacent oppositely charged layer and metal nanoparticle formation. As support, not only flat surfaces but also various nanoparticles can be used. Such supports allow one to obtain core–shell materials or hollow spheres after further treatment (for example, dissolution of a core at certain conditions) [10, 11]. Each individual polyelectrolyte layer can serve as a nanoreactor.



(CdPSS) **(P4VP)** **(PEI)**

Reference [12] describes preparation of CdS nanoparticles in the material which originated from alternate deposition of cadmium polystyrenesulfonate (PSS) and protonated polyvinylpyridine [P4(2)VP] on quartz and other supports modified by protonated polyethyleneimine (PEI) (to charge the surface). Afterward the multilayer material was treated with H_2S which resulted in CdS nanoparticle formation within a polysulfonate layer. In some cases nanoparticles were prepared in the surfactant micelles to ensure better control over particle growth and then mixed with PE to form multilayer films with nanoparticles [13]. In the latter case, particles were formed outside of the nanostructured polymer layer (in surfactant micelles) so it falls out of the topic of this chapter. At the same time, taking into consideration recent discussion on the presence of a broad interface in polyelectrolyte multilayer films [14] (unlike amphiphilic block copolymers), one could consider incorporation of preformed nanoparticles being a viable approach to control particle size distribution. In the case of a broad interface, "nanoreactor" size strongly varies which can cause broader particle size distribution.

2.1.3. Polyelectrolyte Gel–Surfactant Complexes

Polyelectrolyte gel–surfactant complexes (PGSCs) are highly ordered systems which were extensively studied by theorists and experimentalists [15–18]. In these systems, surfactant molecules self-assemble and their ordering even exceeds the ordering of surfactant in the solutions [19]. As PGSCs contain a number of ions in well-ordered surfactant areas inside a weakly cross-linked gel, their replacement with ions of interest (transition metal ions) should offer the possibility of metal nanoparticle formation in a nanostructured medium. As found, metal ion uptake takes place when metal ions are oppositely charged toward polyelectrolyte charged groups [20–22]. In doing so, the incorporated metal compounds form clusters which are located between charged groups of a polymer chain and surfactant head groups. These clusters behave as nuclei for metal nanoparticle formation during reduction. If a fast reducing agent is used ($NaBH_4$), particles grow completely in an ordered environment. Moreover, as reported in [22], although particle growth decreases the degree of crystalline ordering in PGSCs, the ordering around growing particles even increases, which results in additional control over particle growth and a very narrow particle size distribution (Fig. 2). If reduction is carried out with a sluggish reducing agent ($N_2H_4 \times H_2O$), small clusters and ions migrate in unstructured areas of the gel where uncontrolled particle growth occurs. This results in large particles with a very broad particle size distribution. PE gels without internal nanostructure are also unable to control the particle growth.

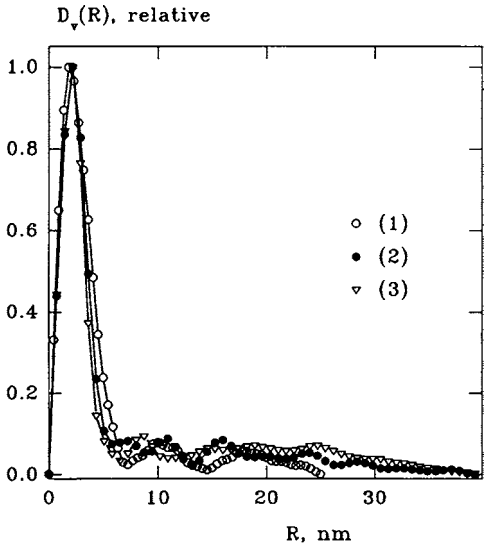


Figure 2. Volume distribution functions for the PDADMACl/sodium dodecylbenzosulphonate (SDBS) gel containing H_2PtCl_6 evaluated from anomalous small angle X-ray scattering (ASAXS) difference curves. (1): compound clusters; (2) and (3): Pt nanoparticles reduced with $NaBH_4$ and $N_2H_4 \cdot H_2O$, respectively. The distributions are normalized to a maximum value of unity. Reprinted with permission from [22], D. I. Svergun et al., *J. Phys. Chem. B* 104, 5242 (2000). © 2000, American Chemical Society.

2.2. Nanoparticles in Polymer Cavities

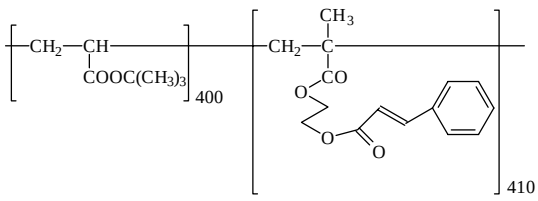
Existence of nanopores or nanocavities of a regular size and shape in polymer matrix creates interfaces between pore volume and polymer walls, thus creating nanostructures in polymers. These pores can be formed in polymers due to special conditions of polymer synthesis or polymer nature [23] and may serve as nanoreactors for particle growth; in this case, a particle size may be restricted by a pore size. This approach is very common for mesoporous solids where metal compounds are incorporated inside the pores and then nanoparticle growth is carried out within the pores [24, 25], but it is rather rare for polymers as the polymers containing regular interpenetrating cavities and pores are also uncommon. A detailed description of template synthesis of various nanoparticles (polymer, metal, metal oxide, etc.) in nanoporous membranes is given in a review article [23]. Here, porous alumina membranes are the most commonly used templates, although polycarbonate porous membranes were also described.

Two types of cavities will be considered here: (i) with and (ii) without any functional groups which might interact with metal compounds.

2.2.1. Nanoparticles in Nanoporous Polymers

The progress in nanoporous polymer synthesis was strongly inspired by the need for polymers with low dielectric constant [26] although in this case closed pores are required. Commonly this is achieved by templating over block copolymers with a thermally labile block. Another use of porous polymer membranes with open pores is metal deposition which is well described in [27].

An interesting example of a porous polymeric system is polymer microspheres containing PAA-lined channels which are most likely continuous [28]. These microspheres were prepared by ultraviolet (UV) crosslinking of poly(*t*-butyl acrylate)-*block*-poly(2-cinnamoyloxyethyl methacrylate) (P*t*BA-*b*-PCEMA) block copolymer micelles followed by hydrolysis.



(P*t*BA-*b*-PCEMA)

The microspheres were used for Pd nanoparticle formation that was performed by incorporation of the microspheres in the $Pd(NO_3)_2$ aqueous solution. Pd^{2+} ion uptake was 1 Pd ion per two carboxylic groups. To increase Pd loading, further Pd deposition during electroless plating has been used. This allowed an increase of Pd content to 27 and 63 wt% (from 13 to 14 wt%) which resulted in an increase of the Pd particle size. As the pores in the microspheres were comparatively large (~30 nm), they might provide restriction of the particle growth only at very high loading. In doing so, particles are getting so large that the advantages of the metal confinement are nearly lost (at least for catalytic applications).

Hyper-cross-linked polystrene (HPS) [29, 30] is another porous material of choice. This is a unique polymer network where pores form spontaneously during polymer synthesis. Due to its high cross-link density, which can exceed 100%, HPS consists of nanosized rigid cavities of comparable size in the 2–3 nm range. It is readily produced by chemically incorporating methylene groups between adjacent phenyl rings in dissolved polystyrene homopolymer or gelled poly(styrene-*r*-divinylbenzene) copolymer in the presence of ethylene dichloride. A unique feature of HPS is its ability to swell in a wide variety of different solvents, even thermodynamically poor ones (e.g., water). Such versatility greatly facilitates incorporation of various organometallic compounds into the nanostructured HPS matrix. Though HPS was known for absorption of heavy metal ions from solutions [31], the first use of HPS as a medium for metal nanoparticle formation was reported in 1999 [32]. Cobalt nanoparticles have been selected here due to the established [33] correlation between size and shape of Co particles and their magnetic properties, that is, their ferromagnetic resonance (FMR) spectrum characteristics. Particles measuring less than 1 nm in diameter are nonmagnetic, while those with diameters from 1 to 10 nm are superparamagnetic. Larger Co particles behave as ferromagnetic materials. Moreover, the width of the FMR signal (ΔH) for spherical particles depends on the size of the Co nanoparticles [33, 34], and the position corresponding to the zero signal (H_0) contains information regarding the shape of the Co nanoparticles. This allows one to determine the general size and shape characteristics of Co nanoparticles solely from FMR spectra.

Impregnation of HPS by either $Co_2(CO)_8$ in 2-propanol or the $[Co(DMF)_6]^{2+}[Co(CO)_4]_2^-$ complex in dimethylformamide (DMF), followed by thermolysis at 200 °C, results in the formation of discrete Co nanoparticles. The concentration and characteristics of such nanoparticles were investigated by X-ray fluorescence spectroscopy, FMR spectroscopy, and transmission electron microscopy (TEM). The FMR data confirmed the formation of spherical nanoparticles. At Co concentrations of 2–8 wt% the magnitude of the FMR linewidth reveals that the mean Co nanoparticle diameter is about 2 nm, which agrees well with the mean particle diameter discerned by TEM. An increase in Co content higher than 8 wt% is accompanied by an increase in mean particle diameter due to an increase in the population of large Co nanoparticles up to 15 nm across. Regulated nanoparticle growth over a wide range of Co concentrations is attributed to nanoscale HPS cavities, which serve to physically restrict the size of growing particles. It is noteworthy to mention that the HPS employed does not contain functional groups that could specifically interact with growing metal nanoparticles and stabilize them. In fact, the nanoscale cavities in HPS are anticipated to be highly interconnected and, therefore, should *not* hinder the migration of Co atoms or even small Co clusters. A steric limitation arising from the predominance of phenyl rings represents the most probable reason for controlled Co nanoparticle growth in HPS; phenyl rings may interact with metal surfaces and nonspecifically stabilize metal nanoparticles. If the Co concentration in a HPS-Co composite exceeds a saturation limit (at Co loadings in excess of 10 wt%), a few large nanoparticles grow. Their size might (i) correlate with the statistics of the pore size distribution of the HPS matrix or (ii) identify a metal-induced rearrangement of the HPS network.

As was shown by the example of Co particle formation, HPS is a robust and convenient nanostructured polymer matrix which might be broadly used for growth of catalytically active nanoparticles. Following this path, the authors [35] reported formation of Pt nanoparticles in HPS. Impregnation of HPS with tetrahydrofuran (THF) or methanol (ML) solutions containing platinic acid (H_2PtCl_6) results in the formation of Pt(II) complexes within the nanocavities of HPS. Subsequent reduction of the complexes by H_2 yields stable Pt nanoparticles with a mean diameter of 1.3 nm in THF and 1.4 nm in ML. As discussed above, for Co-containing HPS, the mean particle size was about 2 nm. This discrepancy in particle size can be explained by the different mechanisms of Co and Pt particle formation. As Co particles were formed at 200 °C, Co clusters could easily migrate between the HPS pores forming Co nanoparticles with the sizes matching the pore size. For Pt nanoparticles, H_2 reduction occurs at room temperature and migration of Pt clusters or atoms or Pt precursor molecules from pore to pore hardly occurs. This results in the formation of Pt particles with a size matching the size of Pt precursor (salt or complex) filling the pores [35]. Thus, for Co nanoparticle formation, HPS controls the particle size due to physical limitations of their growth by the pore size (2 nm). In the case of Pt nanoparticles, pores limit the amount of precursor filling the single pore so during reduction the precursor particle shrinks (density increases from Pt complex to Pt metal) to Pt nanoparticle (1.3 nm in diameter). In both ways, HPS provides strong control over nanoparticle growth. The simplicity of this approach makes this polymer system very attractive for broader use (semiconductor and other catalytic particle formation). The limitation of this system lies in the cross-linked nature of HPS, which does not allow film and coating formation. This limitation is of no concern from the viewpoint of catalytic materials (moreover, it can be used without support in heterogeneous catalysis) but strongly limits its prospects for HPS use for optical or magnetic materials. Another drawback of metallated HPS is an easy washing of the metal nanoparticles out of polymer in media where PS swells. So to prevent metal loss, HPS based catalysts should be used only in "poor" (for PS) solvents (water is the best).

2.2.2. Cavities of Nanostructured Polyoctadecylsiloxane

If the nanoporous polymer contains functional groups inside the pores, incorporation of metal compounds is facilitated and metal particle stability is enhanced. As an example, polyoctadecylsiloxane (PODS) containing nanocavities formed by a bisiloxane layer will be discussed here [36, 37]. This polymer is obtained by hydrolytic polycondensation of octadecyltrichlorosilane in water. As demonstrated in [36], the polymeric condensate consists of highly uniform, pillared microcrystallites in which the inorganic siloxy backbones form periodic layers, each containing a monomolecular layer of intercalated water, separated by crystalline assemblies of alkyl chains. These siloxy bilayers form nanocavities containing silanol groups. Being placed in an aqueous solution of metal salt, this polymer system absorbs salt into its

hydrophilic layer where, due to subsequent reduction, regularly spaced metal nanoparticles were formed [38, 39]. The incorporation of metal compounds in the nanostructured polymer containing siloxy bilayers is strongly determined by the type of metal compound. Planar ions ($PdCl_4^{2-}$, $PtCl_4^{2-}$) presumably driven by entropy are found to easily penetrate within thin layers of 2–3 nm in size. The entropy gain is provided by the hydrogen bond formation between the chlorine atoms of metal anions and the silanol groups so that each $PtCl_4^{2-}$ or $PdCl_4^{2-}$ ion replaces several water molecules. Subsequent reduction results in metal nanoparticles located in the siloxy bilayers (Fig. 3).

As shown by ASAXS and TEM, the sizes of the metal nanoparticles formed in the restricted environment of a siloxy bilayer do not depend on the reducing agent type, the metal compound loading, and the metal compound type. This points to a "cage" effect whereby the size of the cavity limits the particle growth. The increase in electron density upon nanoparticle formation permits direct visualization of these bilayers, as well as the individual nanoparticles residing within them (Fig. 4). In Figure 4 one can see that particles are located solely in siloxy bilayers. These nanoparticles measure about 1–2 nm in diameter and possess a narrow size distribution due presumably to volume availability within the ordered bilayers of PODS. Thus, the cavity size of the nanostructured polymer controls the nanoparticle growth. A modification of PODS with surfactants (cetylpyridinium chloride) or other long chain functionalized molecules allows immobilization of the metal ions both in the siloxy bilayer and in the hydrophobic part of the polymer. In this case, the reduction conditions still do not influence the particles growing in the siloxy bilayer but affect those located in the hydrophobic layers (no "cage" effect for the latter). Since the PODS nanostructure remains unaffected by the interaction with metal ions and reducing agents, this polymer seems to be robust enough to produce not only highly stable metal nanoparticles but also a nanostructured polymer containing regularly spaced nanoparticles. On the other hand, the amount of metal species being absorbed by the PODS cavities is very limited so the overall metal content is low. Another downside of this material is its insolubility in any media, preventing film formation.

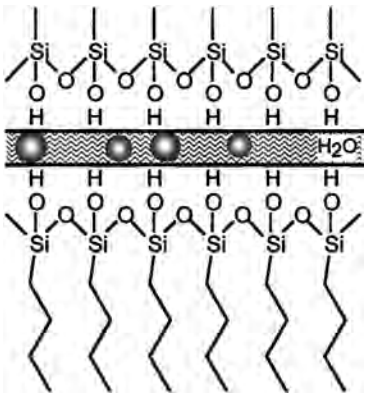


Figure 3. Schematic representation of metal nanoparticle-containing PODS. Reprinted with permission from [38], L. M. Bronstein et al., *Langmuir* 16, 8221 (2000). © 2000, American Chemical Society.

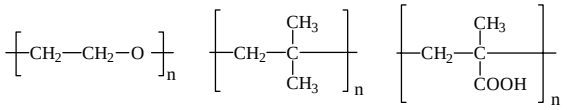
Figure 4. Representative TEM micrograph of PODS-Pd including Pd nanoparticles derived from Na_2PdCl_4 and embedded within the siloxy bilayers of the PODS lamellar nanostructure. The 3 × enlargement more clearly displays the nanoparticles within their host medium. Reprinted with permission from [38], L. M. Bronstein et al., *Langmuir* 16, 8221 (2000). © 2000, American Chemical Society.

2.3. Solutions of Nanostructured Polymers

If nanostructured polymers are not cross-linked, they form stable colloidal solutions. These colloids (block copolymer micelles, dendrimers, microgels, etc.) play the role of nanoreactors for metal, metal oxide, or semiconductor nanoparticle formation.

2.3.1. Nanoparticles in Block Copolymer Micelles

Block Copolymer Micelle Cores Amphiphilic block copolymers form micelles in selective solvents (a good solvent only for one block); yet the size and shape of micelles depend on the block chemical structure, molecular weight of each block, and solvent type [1, 3]. If the micelle core consists of the block bearing functional groups (able to react with metal compounds, forming complexes or salts), such a core can be filled with a corresponding metal compound (by incorporating the metal compound in the block copolymer solution) and further can serve as a nanoreactor for nanoparticle formation. At the same time, the micelle core can be treated like a quasi-solid state, as the core-forming block is not soluble in a selective solvent. As such "functional" blocks, one may use polyvinylpyridines (P2(4)VP), polymethacrylic (PMAA) and PAA acids, PB, PI, and some others. Micelle corona should be formed by a block containing no functional groups but providing solubility and micelle stability in the respective solvent. These can be PS, poly(ethylene oxide) (PEO), polyisobutylene (PIB), etc.

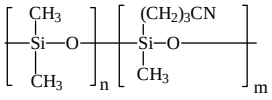


(PEO) **(PIB)** **(PMAA)**

Synthesis of metal or semiconductor nanoparticles in the cores of amphiphilic block copolymer micelles was reported almost simultaneously by several research groups [5, 40–43]. The Au, Pd, Pt, and other particle formation in block

copolymer micelles derived from PS-*b*-P4VP block copolymers showed that nanoparticle morphology strongly depends on the type of reducing agent. When a sluggish reducing agent is used, one nanoparticle per micelle ("cherrylike" morphology) can be formed if there is no exchange between micelles (for example, the micelle is cross-linked) [44]. Fast reduction leads to formation of many small particles per micelle ("raspberrylike" morphology) which is considered to be preferable for catalytic applications (Fig. 5) [45]. Using block copolymer micelle cores as nanoreactors allows synthesis of mono- and bimetallic nanoparticles; yet bimetallic particle morphology depends on a metal pair [46].

Co nanoparticles of different size and shape can be prepared either by incorporation of $CoCl_2$ in the PS-*b*-P2VP micelles followed by reduction or by thermal decomposition of $Co_2(CO)_8$ species embedded in the micelle cores [47]. Stable suspensions of superparamagnetic cobalt nanoparticles were prepared in poly(dimethylsiloxane) (PDMS) carrier fluids in the presence of poly[dimethylsiloxane-*block*-(3-cyanopropyl)methylsiloxane-*block*-dimethylsiloxane] (PDMS-*b*-PCPMS-*b*-PDMS) triblock copolymers as steric stabilizers [48].



(PDMS-*b*-PCPMS)

These copolymers formed micelles in toluene and served as nanoreactors for thermal decomposition of the $Co_2(CO)_8$ precursor. The nitrile groups on the PCPMS central blocks are thought to adsorb onto the particle surface, while the PDMS end-blocks protrude into the reaction medium to provide steric stability. The particle size can be controlled by adjusting the cobalt to copolymer ratio. TEM shows nonaggregated cobalt nanoparticles with a narrow size distribution and the particles are evenly surrounded with copolymer covering.

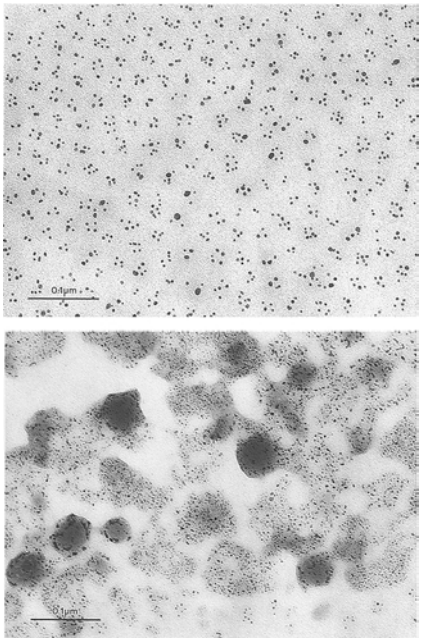
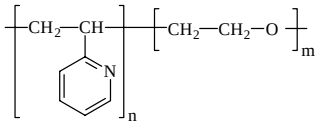


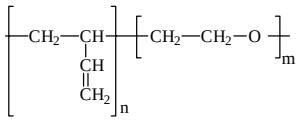
Figure 5. Electron micrographs of Pd colloids synthesized in PS-*b*-P4VP block copolymers via reduction with hydrazine (top) and $NaBH_4$ (bottom). Reprinted with permission from [45], M. V. Seregina et al., *Chem. Mater.* 8, 923 (1997). © 1997, American Chemical Society.

Formation of iron oxide particles in cross-linked block copolymer micelles is described in [49]. A polyisoprene-*block*-poly(2-cinnamoylethyl methacrylate)-*block*-poly(*tert*-butyl acrylate), PI-*b*-PCEMA-*b*-P*t*BA, forms spherical micelles in THF/hexane mixture with 65% volume fraction of the latter. The micelles consist of a PI corona, a solvent-insoluble PCEMA shell, and P*t*BA core. Their structure is locked in by photo-cross-linking the PCEMA shell to yield nanospheres. Similar to core cross-linking, this approach prevents exchange between micelles. The nanospheres were made water-dispersible by hydroxylating the PI double bonds. The core was made compatible with inorganic species by removing the *tert*-butyl groups of P*t*BA. The possibility of using such nanospheres as nanoreactors for inorganic nanoparticle preparation was demonstrated by incorporating iron salt and formation of iron oxide magnetic particles in the cores. However, this approach seems to be far too complex for magnetic particle formation as more robust and simple approaches can be used.

All the above examples describe nanoparticle formation in block copolymer micelles in organic solvents. Such examples are numerous, as many amphiphilic block copolymers form micelles with functionalized core in the organic medium. The choice of block copolymers for aqueous medium is very limited and metal particle formation is normally more complicated as the pH of the medium should be taken into consideration. A few examples of such block copolymers include P2VP-*b*-PEO and PB-*b*-PEO [50, 51]; yet the former block copolymer micellization depends on pH [52]. At pH below 5 P2VP-*b*-PEO becomes molecularly soluble in water. Nevertheless, lowering the pH of the P2VP-*b*-PEO micellar solution after incorporation of metal compounds or metal nanoparticle formation does not result in micelle decomposition, though micelles become less dense. In the case of PB-*b*-PEO, micelles formed in water are very dense, so they successfully fulfill two roles: they serve as nanoreactors for Pd, Pt, and Rh nanoparticle formation and as metal-particle-containing templates for mesoporous silica casting.



(P2VP-*b*-PEO)



(PB-*b*-PEO)

If the P2VP block is a middle block in PS-*b*-P2VP-*b*-PEO triblock copolymer, the "layered" well-defined micelles are formed in water with the PS core, P2VP shell, and PEO corona [53]. Here gold nanoparticle formation was carried out in the P2VP shell serving as a nanoreactor. As the shell is formed by the pH-sensitive P2VP block, according to the author's opinion this system can be useful for encapsulation and/or release of active species. However, one should remember that after metal particle formation this block loses its ability to dissolve at low pH [50]. So this property

can be used if no nanoparticles providing quasi-cross-linking are formed in the P2VP shell.

An original approach to form spherical assemblies of CdS-containing block copolymer reverse micelles in aqueous solution was reported in [54]. These stable assemblies were formed by slow addition of water to mixtures of the reverse micelles formed by PS-*b*-PAA and single PS-*b*-PAA chains. The structures are large compound micelles (LCMs) with quantum-confined CdS nanoparticles dispersed throughout a spherical PS matrix, which is stabilized in water by a layer of solubilized hydrophilic chains. The size of the CdS particles (approximately 3 nm) is determined by the ionic block length of the block copolymer forming the reverse micelle. LCM formation is dependent on the amount of added stabilizing copolymer. This method allows one to transfer the CdS nanoparticles formed in the micelle cores in organic medium to aqueous medium without loss of stability and particle aggregation.

Block Copolymer Micelle Coronas Nanoparticle synthesis can be carried out in the corona of amphiphilic block copolymer micelles. However, if the corona is functionalized, addition of metal salt can result in immediate formation of large aggregates and their precipitation so such a synthesis can be performed only in very diluted solutions. If the corona does not contain groups able to coordinate with metal compounds, particle stabilization is realized solely due to hydrophobic interactions with the hydrophobic core. This scenario was followed when synthesis of Pd, Pt, Ag, and Au nanoparticles was performed in aqueous solutions of PS-*b*-PEO and PS-*b*-PMAA by reduction of the corresponding salts in block copolymer solutions [55, 56]. As reported, this results in the formation of nanometer sized colloidal particles. As stabilization of these particles is provided due to hydrophobic interactions with the PS core, the stability of such systems cannot be satisfactory. At the same time, accessibility of particles in the micelle coronas can be favorable from the viewpoint of catalytic properties.

To improve stabilization in the micelle coronas, the authors [57–59] suggested using hybrid micelles consisting of PS-*b*-PEO and surfactants. It was expected that surfactant hydrophobic tails should be embedded in the PS core while surfactant head groups will be located on the micelle core surface or in its vicinity. As shown in Figure 6, exchange of surfactant counterions for ions of interest (Pt, Pd, or Rh) ought to allow saturation of the core with the given ions.

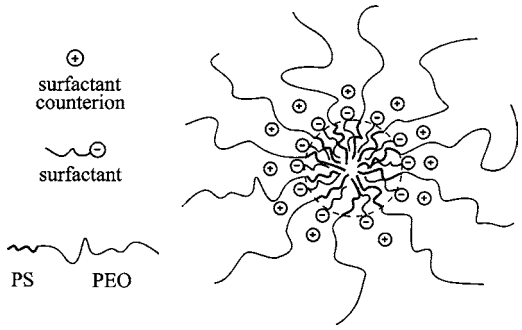


Figure 6. Schematic image of the hybrid PS-*b*-PEO/SDS micelle. Reprinted with permission from [58], L. M. Bronstein et al., *J. Colloid Interface Sci.* 230, 140 (2000). © 2000, Academic Press.

Dynamic light scattering (DLS) and sedimentation in an ultracentrifuge showed that incorporation of positively or negatively charged surfactants results in increase of size and weight of micelles and micellar clusters up to a certain surfactant concentration (which is different for different surfactants). Further increase of surfactant loading (as a rule, above critical micelle concentration for surfactants) results in a moderate decrease of micelle size and weight. By ^{1}H nuclear magnetic resonance, incorporation of surfactant leads to strong increase of mobility of the PS core which proves comicellization of block copolymer molecules and cationic or anionic surfactants. Ion exchange of surfactant counterions in the PS-*b*-PEO/CPC system for $PtCl_6^{2-}$ or $PdCl_4^{2-}$ ions results in saturation of micellar structures with Pt or Pd ions. Subsequent reduction of metal-containing hybrid micellar systems PS-*b*-PEO/CPC/MX_n with $NaBH_4$ or H_2 results in the formation of metal nanoparticles mainly located within the micelles. Morphology and stability of Pd and Pt nanoparticles synthesized in these systems depends on the metal compound loading and the type of a reducing agent. $NaBH_4$ reduction leads to decomposition of micellar clusters and formation of micelles with embedded nanoparticles. These systems display exceptional stability (for years) if metal salt loading does not exceed 1.24×10^{-2} M. Hydrogen reduction results in metal nanoparticle formation both in micelles and micellar clusters so stability of colloidal solutions is provided when metal salt concentration does not exceed 3.36×10^{-3} M. As found, nanoparticle size does not depend on the reducing agent type (contrary to the nanoparticles formed in other block copolymer solutions [42, 50]) but depends on the metal type [57–59]. This could be governed by strong interaction of surfactant head groups with growing nanoparticles.

When the targeted metal can be obtained only as a cation, the hybrid micellar system should include anionic surfactant: sodium dodecylsulfate (SDS) or SDBS. For one, using Rh cations $[Rh(Py)_4Cl_2]^+$, Rh nanoparticles with a diameter of 2–3 nm have been formed in the PS-*b*-PEO/SDS solutions. As seen from the above, incorporation of surfactants in the block copolymer micelles containing no functional groups allows reliable stabilization of metal nanoparticles of 2–6 nm in size. The disadvantage of these systems is the lack of direct methods to vary the particle size.

Micelle Formation via Complexation If both blocks of the corresponding block copolymer are soluble in the particular solvent, nanostructurization can occur due to complexation of the functional groups of the one block with a metal compound. First this phenomenon was described in [60] for the PS-*b*-PB molecular solution in toluene when micellization occurred due to formation of Pd and Pt π-complexes with double bonds. Another example is PS-*b*-P2VP which was molecularly dissolved in a good solvent for both blocks, THF, and subjected to interaction with cadmium ions [61]. It resulted in the formation of aggregates of single micelles called compound micelles. The growth of CdS nanoparticles is confined to the core of single micelles after introduction of hydrogen sulfide gas into the solution. UV-visible spectroscopy, fluorescence spectroscopy, and transmission

electron microscopy were employed to characterize the CdS nanoparticles. UV-visible absorption spectra show that larger nanoparticles are produced at lower 2VP:Cd^{2+} molar ratio, i.e. at higher Cd^{2+} loading. UV-visible absorption spectra and fluorescence spectra both indicate that with decrease of block copolymer concentration in THF, the size of the CdS nanoparticles decreases; that is explained by decrease of the micelle size. Thus, nanoparticle characteristics are dependent on numerous parameters.

If both blocks of a block copolymer are hydrophilic, the block copolymer gives a molecular solution in water. Again, if one of the blocks is inert and another contains functional groups, which are able to interact with metal compounds, their addition leads to micelle formation.

PEO-*b*-PEI was employed for synthesis of metal nanoparticles in aqueous medium using the idea described above [62, 63], as PEI easily interacts with metal compounds due to coordination [64]. This results in aggregation of the PEI block and micelle core formation, while PEO forms a corona and ensures colloidal solubility. Such micelles also provide control over particle nucleation and growth (similar to amphiphilic block copolymer micelles), although micellar and metal particle characteristics are more dependent on the reaction conditions: micelle formation is induced solely by metal compounds and micelle core densities are low. Metal ion reduction leads to metal nanoparticle formation. In so doing, PEO-*b*-PEI micelles remain intact, while changing their hydrodynamic parameters. This gives evidence of strong interaction of the PEI units with metal nanoparticle surface. Metal nanoparticle morphology here can be affected in the same way as in amphiphilic block copolymer micelles by varying the reducing agent type and the metal compound loading along with varying the pH.

The same block copolymer was successfully used for stabilization of CdS nanoparticles [65]. CdS nanoparticles had a well-resolved cubic structure and were monodisperse. The CdS nanoparticles showed a very good resistance against oxidation for months due to their polymer environment. The particle size was controllable in the range between 2 and 4 nm by adjusting the polymer concentration and choice of the solvent.

The ability of PEO-*b*-PMAA to control inorganic morphologies was applied recently to synthesis of silver nanowires [66]. Yet here a block copolymer was used both as a directing and reducing agent. The authors suggest that formation of double hydrophilic block copolymer complex micelles may play an important role in the directed aggregation growth of the silver nanowires. $CdWO_4$ one-dimensional nanorods with diameters of 2.5 nm were obtained in the presence of a similar double-hydrophilic block copolymer: PEO-*b*-PMAA and its partially phosphonated analog [67]. The $CdWO_4$ species formed in block copolymers showed highly increased fluorescence efficiency.

2.3.2. Nanoparticles in Polyelectrolyte Microgels

Microgels are gel spherical particles having diameters in a nanometer range [68, 69]. Cross-linking grants stability of their sizes and properties while their size ensures formation of colloidal solutions. When these microgels are formed by polyelectrolytes, they contain charged groups which are prone to ion exchange. Charged groups also provide solubilization in water. This allows one to consider microgels as nanoreactors for controlling the nanoparticle growth. Gold nanoparticle formation was studied in microgels based on sulfonated PS [70]. Morphologies of nanoparticles formed are strongly determined by the degree of microgel cross-linking. The higher the cross-linking density, the higher the probability of the formation of spherical particles embedded in the microgels. The type of reducing agent is the other key factor. As found, fast reduction ($NaBH_4$) in water results in small gold nanoparticles (4.5 nm) located in microgels. If $NaBH_4$ is added in alkaline solution (0.1 N NaOH), it slows the nucleation down: this results in 7 nm nanoparticles forming long "threads." Yet only 20% of microgels contain nanoparticles. So slow nucleation allows Au clusters and ions migrate out of ordered microgel areas where aggregation easily occurs. High resolution TEM shows that particles grow from one nucleus and microgel environment governs the particle shape. So here microgel provides a "cage" effect although the "cage" is elastic (not rigid as for HPS). In a similar fashion Pd and Pt nanoparticles have been grown in microgels. These metal-particle-containing microgels were used as co-templates (along with amphiphilic PS-*b*-PEO block copolymers) for mesoporous silica formation [71]. Here microgels play a dual role: they are nanoreactors for metal particle formation and pore-forming templates when mesoporous silica is formed.

2.3.3. Nanoparticle Formation in Dendrimers

Metal particle formation in dendrimers can be treated as an intermediate case between particle formation in the nondense block copolymer micelle cores and in the polymer cavities. Most commonly used and studied dendrimers (commercially available) are poly(amidoamines) (PAMAM), the structure of which is presented in [72]. Varying dendrimer size (dendrimer generation) and metal compound loading, one can vary the size of nanoparticles. For the first time dendrimers were used for stabilization of gold and silver nanoparticles with subsequent self-assembling of metal-containing dendrimers on the surface with monolayer formation [73]. Later [74] PAMAM dendrimers with terminal hydroxyl groups were used as templates and stabilizers for controlled synthesis of monodisperse catalytically active nanoparticles. This was achieved by sorption of metal ions inside the dendrimers. The driving force for metal species incorporation is formation of metal complexes with amino groups. Subsequent reduction results in the formation of metal nanoparticles encapsulated in dendrimers; yet the nanoparticle size is exactly determined by the amount of metal atoms loaded in the dendrimer. Deficiency of metal exchange between dendrimer molecules provides very narrow particle size distribution and subtle control over nanoparticle growth. These metal-containing dendrimers are water-soluble and very stable (no precipitation is observed for months). It is considered that a metal particle is formed in the cavity of the dendrimer. This location provides a certain stabilization of the particles, although a "cage" effect (restriction of the particle growth by the cavity size) is not realized. On the other hand, the presence of functional groups (amino groups) allows additional stabilization

of nanoparticles. Metal particle formation resembles the growth of nanoparticles in the cross-linked cores of block copolymer micelles when exchange between micelles does not occur. If PAMAM contains surface (terminal) amino groups, exchange between dendrimer molecules results in a situation when gold particle growth depends on the ratio of amino group/metal compound [75]. As might be expected, it also resulted in a broader particle size distribution than for PAMAM with terminal hydroxyl groups.

Dendrimer generation is an important factor influencing the particle characteristics [76]. Formation of gold colloids upon reduction of a gold salt precursor in charged PAMAM served as a model reaction to study the influence of reaction conditions and dendrimer generation on the resulting nanostructures [77]. Characterizations by TEM, small-angle neutron scattering, and SAXS show that the gold particles are formed inside the dendrimer and located offset from the center (probably in a dendrimer cavity) (Fig. 7). Lower generation dendrimers aggregate while stabilizing metal nanoparticle growth. Dendrimers of generation 6–9 can template one gold nanoparticle per dendrimer molecule, the size of which is well controlled by the number of gold atoms added per dendrimer. These data well agree with those described in [74]. For generation 10, multiple smaller gold particles per dendrimer were observed and dendrimers aggregated (Fig. 8). In [78] new dendrimers, poly(propyleneimine) dendrimers with stearyl end groups, combining both hydrophilic and hydrophobic moieties were used for metal particle formation. These dendrimers form inverse micelles in toluene. While the initial dendrimers have a spherical structure with a collapsed core, solubilization of metal salt hydrate leads to the formation of cylindrical multidendrimer structures with swollen, metal-salt-filled dendrimer cores. When the gold salt inside the dendrimers is reduced to form colloidal particles, the cylindrical structure breaks up and spherical nanoparticles are formed. Yet the particle sizes are larger than would be expected if the gold salt loading of one dendrimer formed one particle, indicating that the ions from several dendrimers are combined. In the author's opinion, these dendrimer-stabilized metal colloids represent hybrid structures that have potential importance as selective catalysts and in the formation of composite materials with special optical properties. On the other hand, the nanoreactors formed by aggregation of dendrimers with formation of dendrimer micelles are less defined and more complex than simple block copolymer micelles. So advantages of such amphiphilic dendrimers are not evident.

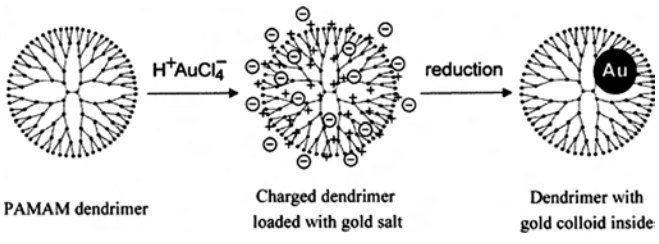


Figure 7. Dendrimer nanotemplating in aqueous solution. In a first step, the dendrimer is loaded with a precursor salt ($H^+AuCl_4^-$), resulting in a charged dendrimer with the precursor as counterions. In a second step, the chemical reduction is performed which yields a colloid inside the dendrimer. Reprinted with permission from [77], F. Gröhn et al., *Macromolecules* 33, 6042 (2000). © 2000, American Chemical Society.

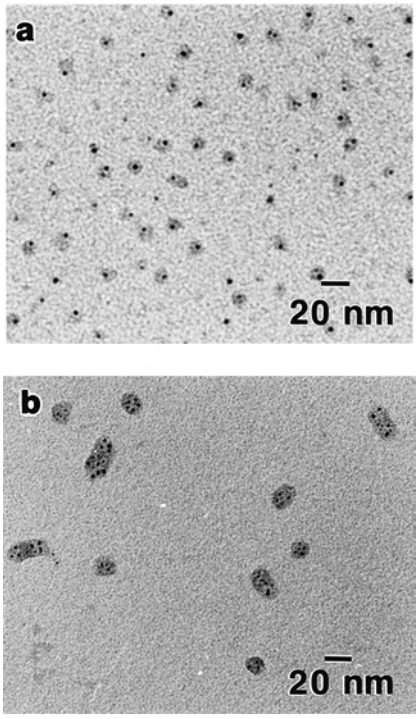


Figure 8. (a) TEM of gold containing G8 PAMAM dendrimer obtained for 1:1 loading and slow reduction. (b) TEM of gold containing G10 PAMAM dendrimer obtained for 1:1 loading and slow reduction. (In both cases the dendrimer has been stained with phosphotungstic acid and appears gray while nanoparticle appears black.) Reprinted with permission from [77], F. Gröhn et al., *Macromolecules* 33, 6042 (2000). © 2000, American Chemical Society.

Along with metal particles, a number of semiconductor particles—CdS [79–81], CdSe [82], and complex core–shell CdSe/ZnS [80]—were successfully prepared in dendrimers. As found in [80], the absorptive and emissive properties of the CdS/PAMAM (hydroxylated) systems are a function of the generation of the dendrimer which is related to the dependence of nanoparticle size on dendrimer generation and the dependence of optical properties on semiconductor nanoparticle size. Additionally, the numerous reactive groups on the surface of the dendrimer can be used to make the composites soluble in essentially any solvent, including water. The luminescence properties of these and other semiconductor nanoparticle/dendrimer nanocomposites will be described below.

3. CHARACTERIZATION

Characterizing nanostructured materials, one faces two challenging tasks: to access nanoparticle characteristics and to get information on nanostructured polymers. Fulfilling these tasks requires using a combination of several modern techniques. The choice of the method is mainly determined by the condition of a polymer matrix (solid state or solution). The following paragraphs give a short description of the major techniques used in characterization of nanostructured polymers with nanoparticles.

Transmission electron microscopy (with magnification 100,000–150,000) is a powerful tool to characterize nanostructured materials with nanoparticles as it allows accessing particle characteristics (size, particle size distribution, particle location) as well as polymer material parameters. Using designated software to process TEM images allows estimation of nanoparticle size and particle size distribution. If one deals with solutions of block copolymer micelles or dendrimers, the TEM sample can be prepared by simply placing a drop of the diluted solution on the electron grid. Micelle or dendrimer size can be obtained from the TEM micrographs along with particle characteristics [42, 43, 74]. Solid materials are normally embedded in the epoxy resin which is then cross-sectioned to allow one to prepare a thin slice. In this case TEM examination allows one to discern the microphase size and shape either in block copolymer bulk or in other solid biphasic materials [7, 38]. The nanoparticles or metal salts in these microphases play the role of staining agents, improving the visualization of the material structure. As seen from the sections above, this method is very commonly used.

One of the key advantages of scanning TEM compared to TEM is that the beam damage is limited to the scanned area since only part of the image is exposed to electrons. This feature can be especially valuable for polymeric systems. Other advantages are high resolution of particle imaging and the possibility of using subnanometer probe and high angle annular dark field (Z-contrast) imaging [83]. Despite these advantages this method is not yet widely used for studying the nanostructured polymers with nanoparticles.

High resolution TEM allows one to obtain more detailed information on nanoparticle size and structure [84, 85]. Along with that, distinguishing the shape of nanocrystals permits conclusions on the nucleation and growth mechanism [70].

Powder X-ray diffraction (XRD) is the convenient method to obtain a mean particle size of crystalline particles [45, 86]. Normally, the corresponding Bragg peaks appearing in the XRD profile allow quantitative assessment of a mean particle size, but no reliable information can be obtained on particle size distribution. Besides, this method cannot be applied to amorphous particles.

Small angle X-ray scattering can be used to characterize the nanoparticles within the nanostructured polymers as well as the nanostructured materials [20, 78]. It allows one to obtain particle size and particle size distribution if the ordered polymer structure does not obscure the assessment of the particle characteristics and does not change polymer ordering during incorporation of metal nanoparticles. In the latter case, ASAXS can be used [22, 39]. Here, the polymer matrix structure does not influence the scattering of the particles as the difference scattering (obtained via subtraction of the scattering curves at the different energies near the adsorption edge of the corresponding metal) is used. Both SAXS and ASAXS can be used for solids and solutions.

Extended X-ray absorption fine structure (EXAFS) and X-ray absorption near-edge structure (XANES) spectroscopies are other methods of choice to characterize nanoparticles and their environments. EXAFS and XANES measurements on the reduced poly(propylene imine) dendrimers with diaminobutane core containing Cu nanoparticles provided additional information on the possible encapsulation of the Cu nanoclusters by the dendrimers and corroborated TEM data [87]. Though these methods are becoming widely used for other types of nanostructured systems (for example, for mesoporous materials with nanoparticles [88]), only a few nanostructured polymers were analyzed in this way. The probable cause of such limited use is the complexity of these methods especially when they are applied to complex nanostructured polymer–nanoparticle systems.

Optical methods (UV absorption and luminescence) are widely used for characterization of the size and size distribution of semiconductor (quantum dots) and some metal (Au, Ag, Cu, etc.) or metal oxide nanoparticles as these parameters determine the optical response [89]. At the same time these methods allow one to characterize the optical properties of nanoparticles [80].

The above methods are designed mainly for particle characterization though TEM and SAXS are also useful for nanostructured polymer description. Other valuable methods for characterization of solutions of nanostructured polymers are DLS [80, 90], static light scattering [91], and sedimentation in an ultracentrifuge [57, 63]. All these methods allow one to obtain sizes of block copolymer micelles, microgels, and dendrimers. Sedimentation in an ultracentrifuge also allows one to distinguish between "empty" polymer nanostructures and those filled with nanoparticles [70].

4. PROPERTIES AND APPLICATIONS

The type of particles embedded in the polymeric materials mainly determines the properties of these materials and their feasible application. Unique properties of nanoparticles can be explained by their huge surface areas and high fraction of surface atoms. This fact is responsible for electron transfer which makes electronic and other properties dependent on the particle size. Both surface area and changed electronic properties are important in catalysis. Electron and spin confinement determines unique optical and magnetic properties of nanoparticles. As discussed in the Section 1, nanoparticles impart their properties to polymeric systems where nanoparticles were grown. In turn, polymers modify the particle surface, often dramatically changing the properties of nanoparticles. Altogether, chemical structure of the polymeric system (functionality and nanostructure) and particle characteristics basically furnish the interface properties and overall material properties [92].

In the following sections, a brief review of the major properties of nanostructured polymers with nanoparticles is presented.

4.1. Catalytic Properties

A majority of noble and transition metal nanoparticles are catalytically active in a number of reactions. Most of the articles dealing with catalytic properties of nanoparticle-containing nanostructured polymers present studies on catalytic hydrogenation of various substrates, as this is an easily available test of catalytic activity. For example, hydroxyl-terminated PAMAM dendrimers containing Pd

nanoparticles showed high catalytic activity in alkene hydrogenation in water [93]. Moreover, the catalytic activity might be controlled by changing the dendrimer size (generation). In the author's opinion, the dendrimer works as a "nanofilter" with a controllable mesh size.

Block copolymer (PS-*b*-P4VP) micelles in organic solvents bearing mono- or bimetallic nanoparticles displayed a high selectivity and activity in hydrogenation of dehydrolinalool to linalool (fragrant substance) [46]. The authors showed that modification of Pd nanoparticles with a second metal alters the electronic properties and surface geometry of the nanoparticles and, as a consequence, catalytic activity. Very high selectivity was provided due to modification of nanoparticle surface with pyridine units of the core-forming block.

Pd nanoparticles stabilized by the above block copolymers were also employed for Heck reaction: cross-coupling of alkylhalides and alkenes [94]. These systems showed reactivity similar to traditionally used Pd complexes but with much higher stability.

Pt nanoparticles formed in HPS [35] were successfully used for direct catalytic oxidation of L-sorbose to 2-keto-L-gulonic acid (vitamin C precursor). Here the catalytic activity and selectivity strongly depend on both the particle size and particle composition. The highest selectivity was obtained for the complex $Pt^0/Pt^{II}/Pt^{IV}$ nanoparticles with a mean diameter of about 1.8 nm obtained *in-situ* during the induction period of catalytic reaction.

So, in general, catalytic properties of such systems depend on the particle size, particle composition, and the type of surrounding polymer system.

4.2. Magnetic Properties

Magnetic properties of particles strongly depend on the particle size: with decrease of the particle diameter, the nature of magnetism changes: ferromagnetic particles become superparamagnetic (do not show hysteresis) and these properties are temperature-dependent. If magnetic particles are formed inside a polymer matrix, magnetic nanocomposites have unique properties and can be used for development of new magnetic and magnetooptical devices [95]. Magnetic measurements are often used for materials containing magnetically active particles to elucidate their composition and structure. In this way, Co nanoparticle formation in the PS-*b*-P4VP micelles and HPS was followed with FMR [32, 47]. In the PS-*b*-P4VP micelles Co nanoparticles were generated by two methods: by reduction of micelles loaded with $CoCl_2$ or by thermal decomposition of $Co_2(CO)_8$ in micellar solutions of these block copolymers [47]. In both cases Co nanoparticles are effectively stabilized by the block copolymer matrix and do not aggregate. For $CoCl_2$, the formed particles have a diameter below 1 nm. Thermal treatment of such dried polymers at 200 °C for 2 h leads to spherical particles of 3–5 nm in size. The polymeric hybrid materials containing the latter particles display remarkably high values of magnetization at rather low Co contents in the polymer (i.e., a tenfold increase of the specific magnetization was obtained). $Co_2(CO)_8$ as a precursor results in more complex behavior. The shape and size of the Co nanoparticles formed by thermolysis can be controlled by the ratio of 4-VP/Co. Both superparamagnetic and ferromagnetic materials can be prepared. For ferromagnetic samples, coercive force is in the range 250–475 Oe depending on the Co content and block copolymer characteristics. Co nanoparticles formed in the PS-*b*-P4VP micelles display remarkable stability under air (FMR signal remains unchanged for months); this confirms strong adsorption of pyridine units on the nanoparticle surface.

Co nanoparticles prepared in triblock copolymer, PDMS-*b*-PCPMS-*b*-PDMS (see Section 5.1.1), showed superparamagnetic behavior; however, some degree of surface oxidation was observed over time, resulting in a decrease in magnetic susceptibility [48].

A novel approach to preparation of magnetic polymeric nanoparticles by synthesis of the magnetite core and polymeric shell [using methacrylate and hydroxyethyl methacrylate (HEMA)] in a single inverse microemulsion was reported in [96]. The polymeric shell size was controlled in the range 80–320 nm by varying the size of the microdroplets. The technique of synthesizing a core and a shell in a single microemulsion has improved the particle structure and size distribution and allowed formation of the particles, made of a superparamagnetic magnetite core coated with a polymeric shell of PMA-*co*-PHEMA random copolymer. Magnetite concentration achieved 3.3 wt%. In a modification of this technique reported in [97], magnetite particles coated with PMAA showed ferromagnetic behavior with coercive force of 50 Oe.

4.3. Optical Properties

UV and luminescence spectra of CdS nanoparticles synthesized in G4 PAMAM with hydroxyl and amino groups [98] showed that the dendrimer stabilizes nanoparticles so their optical properties differ from the CdS bulk. Although dendrimer macromolecules filled with nanoparticles are prone to aggregation, storage at low temperature (−10 °C) prevents the aggregation and solutions retain their photoluminescent properties. The dot–dendrimer nanocomposites can be captured in a SiO_2 sol–gel matrix to yield a stable, bright blue-emitting glass [79].

Similar to dendrimers, Huang et al. [99] prepared CdS-containing nanocomposites using branched PEI as polymer stabilizer. These nanocomposites also demonstrated blue emission but with lower quantum yield. The authors believe that high local concentration of nitrogen atoms (in dendrimer or PEI) is responsible for formation of small well-defined particles.

Semiconductor nanoparticles can be doped with a second metal which strongly modifies the optical properties. Reference [100] reports synthesis of ZnS (doped with Mn and Tb) in microphase-segregated films of diblock copolymers containing carboxylic groups in one block. These films demonstrated emission which is characteristic for doping ions.

A remarkable nanocomposite was prepared by incorporation of CdTe nanoparticles in thin films made by layer-by-layer assembly of an anionic polyelectrolyte with poly(*p*-phenylene ethynylene) backbone (aPPE) and poly(allylamine hydrochloride) polycation [101]. Tethering of CdTe was accomplished by using poly(ethyleneglycol)

(PEG) chains with two reactive termini such as *t*-BOC-NH-PEG-COO-NHS. As found, luminescence spectra after incorporation of particles displays a new peak at 600 nm, identical to the emission peak of CdTe dispersions. Yet the excitation spectrum of CdTe-PEG-aPPE registered at 600 nm reveals a strong new band at 400 nm, which matches the absorption band of aPPE LBL films. This indicates that the light quanta emitted by the nanoparticles originate from the light absorption of the polymer film. The average separation distance between the nanoparticles tethered to the LBL films can be changed by altering the dielectric properties of the solvent affecting PEG tether coiling (water/alcohol mixture). CdTe-PEG-aPPE is an example of an organized nanoparticle system with tunable optical coupling.

4.4. Sensor Properties

Chemical sensors are an important area of the nanoparticle application. Normally, to make a viable device, nanoparticles should be located in a thin film. Polymers are good candidates for nanoparticle stabilization as they easily form films, although particle interaction with polymer functional groups can influence negatively the sensitivity of the sensors. CdS/dendrimer nanocomposites with a range of concentrations of the photoluminescent semiconductor nanocluster and the dendrimer were prepared in methanol and 1:9 methanol:water solutions and used for detection of cyanoacrylate ester fumed and unfumed fingerprints on polyethylene and aluminum foil [102].

Vossmeyer et al. [103] showed that the disulfide-functionalized polyphenylene dendrimers are well suited to layer-by-layer fabrication of gold nanoparticle/dendrimer composite films. During film deposition the dendrimers interact with nanoparticles and form homogeneous and mechanically reinforced films with a good control over the film thickness. When using a film as a chemiresistor sensor device, high sensitivity to vapors of toluene and tetracyanoethylene was observed but its sensitivity to humidity was negligible. The short response times, the high sensitivity to volatile organic compounds, and the low sensitivity to humidity make the material a promising candidate for sensor applications. However, complex response characteristics indicate that the overall sensing mechanism involves several processes and more robust systems should be found.

Another interesting possibility for application of nanoparticles formed in nanostructured polymers is nanolithography. Adsorption of block copolymer micelles containing Au nanoparticles on different substrates (GaAs- or InP-based semiconductor wafers, glass) allows formation of monomicellar films and using these films for nanolithography [104].

5. CONCLUSIONS

The previous sections clearly show that a number of polymeric systems containing nanostructures can provide stabilization of nanoparticles and tailor nanoparticle size, size distribution, and shape. The polymer nanoenvironment also changes the nanoparticle properties via modification of the nanoparticle surface with polymer chains. All of the above allow development of a number of sophisticated polymer nanocomposites with tunable properties and promise for important applications. The area of their possible function determines the prospect for the development of each particular system. For heterogeneous catalysis, nanoporous functionalized polymers with nanoparticles seem to be most promising and robust materials in a number of catalytic reactions, especially if pore size and functionality can be tailored. For solution applications, most promising are block copolymers and dendrimers with nanoparticles. Moreover, block copolymer micelles can form films of various thickness (by spin coating) which is especially important for optical, magnetic, and sensor applications. Both block copolymer and dendrimers can be adsorbed on the surface in a regular manner which is promising for nanolithography. At the same time, one can expect a vigorous development of this field for years to come so new polymeric systems better able to control nanoparticle formation and properties will be developed.

GLOSSARY

Nanocrystal Nanosized crystal.

Nanoenvironment Environment in nanometer scale.

Nanoparticle Particle of nanometer (10^{-9} m) size.

Nanoreactor Nanostructured inhomogeneity in the material where nanoparticles are formed (for example, nanopores or block copolymer micelles).

Nanostructure Structural inhomogeneity of nanometer size.

Nanostructurization Formation of nanostructure.

REFERENCES

1. I. U. Hamley, "The Physics of Block Copolymers." Oxford Univ. Press, Oxford, 1998.
2. F. S. Bates and G. H. Fredrickson, *Phys. Today* 52, 32 (1999).
3. S. Förster and M. Antonietti, *Adv. Mater.* 10, 195 (1998).
4. L. M. Bronstein and P. M. Valetsky, *J. Inorg. Organomet. Polym.* 4, 415 (1994).
5. M. Moffitt, L. McMahon, V. Pessel, and A. Eisenberg, *Chem. Mater.* 7, 1185 (1995).
6. L. Bronstein, M. Seregina, P. Valetsky, U. Breiner, V. Abetz, and R. Stadler, *Polym. Bull.* 39, 361 (1997).
7. J. F. Ciebien, R. T. Clay, B. H. Sohn, and R. Cohen, *New J. Chem.* 22, 685 (1998).
8. T. Hashimoto, M. Harada, and N. Sakamoto, *Macromolecules* 32, 6867 (1999).
9. Y. Lvov, G. Decher, and H. Möhwald, *Langmuir* 9, 481 (1993).
10. G. B. Sukhorukov, E. Donath, H. Lichtenfeld, E. Knippel, M. Knippel, A. Budde, and H. Mohwald, *Colloids Surf. A* 137, 253 (1998).
11. R. A. Caruso, A. Susha, and F. Caruso, *Chem. Mater.* 13, 400 (2001).
12. H. Xiong, Z. Zhou, Z. Wang, Z. Xi, and J. Shen, *Supramol. Sci.* 5, 623 (1998).
13. N. A. Kotov, I. Dekany, and J. H. Fendler, *J. Phys. Chem.* 99, 13065 (1995).
14. G. Decher, in "219th ACS National Meeting," San Francisco, CA, 2000.
15. A. R. Khokhlov, S. G. Starodubtsev, and V. V. Vasilevskaya, *Adv. Polym. Sci.* 109, 123 (1993).
16. A. R. Khokhlov, E. Y. Kramarenko, E. E. Makhaeva, and S. G. Starodubtsev, *Macromol. Theory Simul.* 1, 105 (1992).

17. Y. V. Khandurina, A. T. Dembo, V. B. Rogacheva, A. B. Zezin, and V. A. Kabanov, *Vysokomol. Soedin., Ser. A Ser. B* 36, 235 (1994).
18. H. Okuzaki and Y. Osada, *Macromolecules* 28, 380 (1995).
19. F. Yeh, E. L. Sokolov, A. R. Khokhlov, and B. Chu, *J. Am. Chem. Soc.* 118, 6615 (1996).
20. L. M. Bronstein, O. A. Platonova, A. N. Yakunin, I. M. Yanovskaya, P. M. Valetsky, A. T. Dembo, E. E. Makhaeva, A. V. Mironov, and A. R. Khokhlov, *Langmuir* 14, 252 (1998).
21. L. M. Bronstein, O. A. Platonova, A. N. Yakunin, I. M. Yanovskaya, P. M. Valetsky, A. T. Dembo, E. S. Obolonkova, E. E. Makhaeva, A. V. Mironov, and A. R. Khokhlov, *Colloids Surf. A* 147, 221 (1999).
22. D. I. Svergun, E. V. Shtykova, M. B. Kozin, V. V. Volkov, A. T. Dembo, E. V. J. Shtykova, L. M. Bronstein, O. A. Platonova, A. N. Yakunin, P. M. Valetsky, and A. R. Khokhlov, *J. Phys. Chem. B* 104, 5242 (2000).
23. J. C. Hulteen and C. R. Martin, in "Nanoparticles and Nanostructured Films" (J. H. Fendler, Ed.), p. 235. Wiley–VCH, Weinheim, 1998.
24. Y. Plyuto, J.-M. Berquer, C. Jacquiod, and C. Ricolleau, *Chem. Commun.* 1653 (1999).
25. L. M. Bronstein, S. Polarz, B. Smarsly, and M. Antonietti, *Adv. Mater.* 13, 1333 (2001).
26. K. R. Carter, R. A. DiPietro, M. I. Sanchez, and S. A. Swanson, *Chem. Mater.* 13, 213 (2001).
27. M. W. Ellsworth and D. L. Gin, *Polym. News* 24, 331 (1999).
28. Z. Lu, G. Liu, H. Phillips, J. M. Hill, J. Chang, and R. A. Kydd, *Nano Lett.* 1, 683 (2001).
29. V. A. Davankov and M. P. Tsyurupa, *React. Polym.* 13, 27 (1990).
30. M. P. Tsyurupa and V. A. Davankov, *J. Polym. Sci.: Polym. Chem. Ed.* 18, 1399 (1980).
31. R. M. C. Sutton, S. J. Hill, and P. Jones, *J. Chromatogra. A* 789, 389 (1997).
32. S. N. Sidorov, L. M. Bronstein, V. A. Davankov, M. P. Tsyurupa, S. P. Solodovnikov, P. M. Valetsky, E. A. Wilder, and R. J. Spontak, *Chem. Mater.* 11, 3210 (1999).
33. S. V. Vonsovskii, "Magnetism." Nauka, Moscow, 1971.
34. C. P. Bean, J. D. Livingston, and P. S. Rodbell, *Acta Metall.* 5, 682 (1957).
35. S. N. Sidorov, I. V. Volkov, V. A. Davankov, M. P. Tsyurupa, P. M. Valetsky, L. M. Bronstein, R. Karlinsey, J. W. Zwanziger, V. G. Matveeva, E. M. Sulman, N. V. Lakina, E. A. Wilder, and R. J. Spontak, *J. Am. Chem. Soc.* 123 (2001).
36. A. N. Parikh, M. A. Schivley, E. Koo, K. Seshadri, D. Aurentz, K. Mueller, and D. L. Allara, *J. Am. Chem. Soc.* 119, 3135 (1997).
37. W. R. Thompson and J. E. Pemberton, *Langmuir* 11, 1720 (1995).
38. L. M. Bronstein, D. M. Chernyshov, P. M. Valetsky, E. A. Wilder, and R. J. Spontak, *Langmuir* 16, 8221 (2000).
39. D. I. Svergun, M. B. Kozin, P. V. Konarev, E. V. Shtykova, V. V. Volkov, D. M. Chernyshov, P. M. Valetsky, and L. M. Bronstein, *Chem. Mater.* 12, 3552 (2000).
40. M. Antonietti and S. Henz, *Nachr. Chem. Technol. Lab.* 40, 308 (1992).
41. H. Saito, S. Okamura, and K. Ishizu, *Polymer* 33, 1099 (1992).
42. M. Antonietti, E. Wenz, L. Bronstein, and M. Seregina, *Adv. Mater.* 7, 1000 (1995).
43. J. P. Spatz, A. Roescher, and M. Möller, *Adv. Mater.* 8 (1996).
44. L. Bronstein, M. Antonietti, and P. Valetsky, in "Nanoparticles and Nanostructured Films" (J. H. Fendler, Ed.), p. 145. Wiley–VCH, Weinheim, 1998.
45. M. V. Seregina, L. M. Bronstein, O. A. Platonova, D. M. Chernyshov, P. M. Valetsky, J. Hartmann, E. Wenz, and M. Antonietti, *Chem. Mater.* 9, 923 (1997).
46. L. M. Bronstein, D. M. Chernyshov, I. O. Volkov, M. G. Ezernitskaya, P. M. Valetsky, V. G. Matveeva, and E. M. Sulman, *J. Catal.* 196, 302 (2000).
47. O. A. Platonova, L. M. Bronstein, S. P. Solodovnikov, I. M. Yanovskaya, E. S. Obolonkova, P. M. Valetsky, E. Wenz, and M. Antonietti, *Colloid Polym. Sci.* 275, 426 (1997).
48. M. Rutnakornpituk, M. S. Thompson, L. A. Harris, K. E. Farmer, A. R. Esker, J. S. Riffle, J. Connolly, and T. G. St. Pierre, *Polymer* 43, 2337 (2002).
49. R. S. Underhill and G. Liu, *Chem. Mater.* 12, 2082 (2000).
50. L. M. Bronstein, S. N. Sidorov, P. M. Valetsky, J. Hartmann, H. Coelfen, and M. Antonietti, *Langmuir* 15, 6256 (1999).
51. L. Bronstein, E. Kraemer, B. Berton, C. Burger, S. Foerster, and M. Antonietti, *Chem. Mater.* 11, 1402 (1999).
52. T. J. Martin, K. Prochazka, P. Munk, and S. E. Webber, *Macromolecules* 29, 6071 (1996).
53. J.-F. Gohy, N. Willet, S. Varshney, J.-X. Zhang, and R. Jerome, *Angew. Chem. Int. Ed.* 40, 3214 (2001).
54. M. Moffitt, H. Vali, and A. Eisenberg, *Chem. Mater.* 10, 1021 (1998).
55. A. B. R. Mayer, J. E. Mark, and R. E. Morris, *Polym. J.* 30, 197 (1998).
56. A. B. R. Mayer and J. E. Mark, *Colloid Polym. Sci.* 275, 333 (1997).
57. L. M. Bronstein, D. M. Chernyshov, G. I. Timofeeva, L. V. Dubrovina, P. M. Valetsky, E. S. Obolonkova, and A. R. Khokhlov, *Langmuir* 16, 3626 (2000).
58. L. M. Bronstein, D. M. Chernyshov, G. I. Timofeeva, L. V. Dubrovina, P. M. Valetsky, and A. R. Khokhlov, *J. Colloid Interface Sci.* 230, 140 (2000).
59. L. M. Bronstein, D. M. Chernyshov, E. Vorontsov, G. I. Timofeeva, L. V. Dubrovina, P. M. Valetsky, S. Kazakov, and A. R. Khokhlov, *J. Phys. Chem. B* 105, 9077 (2001).
60. L. M. Bronstein, M. V. Seregina, O. A. Platonova, Y. A. Kabachii, D. M. Chernyshov, M. G. Ezernitskaya, L. V. Dubrovina, T. P. Bragina, and P. M. Valetsky, *Macromol. Chem. Phys.* 199, 1357 (1998).
61. H. Zhao, E. P. Douglas, B. S. Harrison, and K. S. Schanze, *Langmuir* 17, 8428 (2001).
62. L. M. Bronstein, S. N. Sidorov, A. Y. Gourkova, P. M. Valetsky, J. Hartmann, M. Breulmann, H. Colfen, and M. Antonietti, *Inorg. Chim. Acta* 280, 348 (1998).
63. S. N. Sidorov, L. M. Bronstein, P. M. Valetsky, J. Hartmann, H. Colfen, H. Schnablegger, and M. Antonietti, *J. Colloid Interface Sci.* 212, 197 (1999).
64. M. Michaelis and A. Henglein, *J. Phys. Chem.* 96, 4719 (1992).
65. L. Qi, H. Coelfen, and M. Antonietti, *Nano Lett.* 1, 61 (2001).
66. D. Zhang, L. Qi, J. Ma, and H. Cheng, *Chem. Mater.* 13, 2753 (2001).
67. S.-H. Yu, M. Antonietti, H. Colfen, and M. Giersig, *Angew. Chem. Int. Ed.* 41, 2356 (2002).
68. M. Antonietti, *Angew. Chem. Int. Ed.* 27, 1743 (1988).
69. M. Antonietti, W. Bremser, and M. Schmidt, *Macromolecules* 23, 3796 (1990).
70. M. Antonietti, F. Gröhn, J. Hartmann, and L. Bronstein, *Angew. Chem. Int. Ed.* 36, 2080 (1997).
71. N. T. Whilton, B. Berton, L. Bronstein, H.-P. Hentze, and M. Antonietti, *Adv. Mater.* 11, 1014 (1999).
72. F. Zeng and S. C. Zimmerman, *Chem. Rev.* 97, 1681 (1997).
73. G. Bar, S. Rubin, R. W. Cutts, T. N. Taylor, and T. A. Zawodzinski, Jr., *Langmuir* 12, 1172 (1996).
74. M. Zhao and R. M. Crooks, *Adv. Mater.* 11, 217 (1999).
75. K. Esumi, A. Suzuki, N. Aihara, K. Usui, and K. Torigoe, *Langmuir* 14, 3157 (1998).
76. R. M. Crooks, B. I. Lemon III, L. Sun, L. K. Yeung, and M. Zhao, *Topics Curr. Chem.* 212, 81 (2001).
77. F. Gröhn, B. J. Bauer, Y. A. Akpalu, C. L. Jackson, and E. J. Amis, *Macromolecules* 33, 6042 (2000).
78. F. Gröhn, B. J. Bauer, and E. J. Amis, *Macromolecules* 34, 6701 (2001).

79. K. Sooklal, J. Huang, C. J. Murphy, L. Hanus, and H. J. Ploehn, *Mater.* 439 (1999).
80. B. I. Lemon and R. M. Crooks, *J. Am. Chem. Soc.* 122, 12886 (2000).
81. J. J. J. M. Donners, R. Hoogenboom, A. P. H. J. Schenning, P. A. van Hal, R. J. M. Nolte, E. W. Meijer, and N. A. J. M. Sommerdijk, *Langmuir* 18, 2571 (2002).
82. F. Seker and A. B. Ellis, *Macromolecules* 33, 582 (2000).
83. D. Ozkaya, J. M. Thomas, D. S. Shephard, T. Maschmeyer, B. F. G. Johnson, G. Sankar, and R. Oldroyd, *Inst. Phys. Conf. Ser.* 153, 403 (1997).
84. B. Liu, H. Li, C. H. Chew, W. Que, Y. L. Lam, C. H. Kam, L. M. Gan, and G. Q. Xu, *Mater. Lett.* 51, 461 (2001).
85. A. A. Patel, F. Wu, J. Z. Zhang, C. L. Torres-Martinez, R. K. Mehra, Y. Yang, and S. H. Risbud, *J. Phys. Chem.* 104, 11598 (2000).
86. D. H. Gray and D. L. Gin, *Chem. Mater.* 10, 1827 (1998).
87. P. N. Floriano, C. O. Noble, J. M. Schoonmaker, E. D. Poliakoff, and R. L. McCarley, *J. Am. Chem. Soc.* 123, 10545 (2001).
88. J. M. Thomas and G. Sankar, *J. Synchrotron Rad.* 8, 55 (2001).
89. A. D. Yoffe, *Adv. Phys.* 50, 1 (2001).
90. J. Zheng, M. S. Stevenson, R. S. Hikida, and P. G. Van Patten, *J. Phys. Chem. B* 106, 1252 (2002).
91. S. Moessmer, J. P. Spatz, M. Moeller, T. Aberle, J. Schmidt, and W. Burchard, *Macromolecules* 33, 4791 (2000).
92. J. H. Fendler and Y. Tian, in "Nanoparticles and Nanostructured Films" (J. H. Fendler, Ed.), p. 429. Wiley–VCH, Weinheim, 1998.
93. M. Zhao and R. M. Crooks, *Angew. Chem. Int. Ed.* 38, 364 (1999).
94. S. Klingelhöfer, W. Heitz, A. Greiner, S. Oestreich, S. Förster, and M. Antonietti, *J. Am. Chem. Soc.* 119, 10116 (1997).
95. J. R. Thomas, *J. Appl. Phys.* 37 (1966).
96. P. A. Dresco, V. S. Zaitsev, R. J. Gambino, and B. Chu, *Langmuir* 15, 1945 (1999).
97. V. S. Zaitsev, D. S. Filimonov, I. A. Presnyakov, R. J. Gambino, and B. Chu, *J. Colloid Interface Sci.* 212, 49 (1999).
98. L. H. Hanus, K. Sooklal, C. J. Murphy, and H. J. Ploehn, *Langmuir* 16, 2621 (2000).
99. J. Huang, K. Sooklal, C. J. Murphy, and H. J. Ploehn, *Chem. Mater.* 11, 3595 (1999).
100. R. S. Kane, R. E. Cohen, and R. Silbey, *Chem. Mater.* 11, 90 (1999).
101. W. Sebastian and N. A. Kotov, *J. Am. Chem. Soc.* 124, 2448 (2002).
102. E. R. Menzel, M. Takatsu, R. H. Murdock, K. Bouldin, and K. H. Cheng, *J. Forens. Sci.* 45, 770 (2000).
103. T. Vossmeyer, B. Guse, I. Besnard, R. E. Bauer, K. Mullen, and A. Yasuda, *Adv. Mater.* 14, 238 (2002).
104. J. P. Spatz, T. Herzog, S. Mössmer, P. Ziemann, and M. Möller, *Adv. Mater.* 11, 149 (1999).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanopinning in High-Temperature Superconductors

J. Horvat

Institute for Superconducting and Electronic Materials, University of Wollongong, New South Wales 2522, Australia

CONTENTS

1. INTRODUCTION

High-temperature superconductivity is one of the most studied phenomena in the last 20 years. This is because of its potential to bring the fascinating world of superconductivity into our everyday life. However, soon after the discovery of high-temperature superconductivity [1–4], it became clear that two main obstacles have to be overcome: vortex pinning and grain connectivity. Improvement of vortex pinning in *high-temperature superconductors* (HTSs) is in the domain of nanotechnology, because of their special intrinsic physical properties. This chapter outlines improvement of vortex pinning in HTSs by introduction of nanosize *pinning centers*.

Because of the wide range of intended readership, the contribution will begin with a simple description of basic principles of superconductivity and vortex pinning. The focus will gradually shift to more specific topics, intended for more advanced readers.

2. SUPERCONDUCTIVITY

We are taught in school that all processes in everyday life are associated with the dissipation of energy. For example, driving a car, more energy is spent than required for mere movement of the car. This is because some of the energy obtained by burning the gasoline is dissipated into the environment through imperfect efficiency of the engine, friction of wheels, and air resistance to the car. Similarly, transport of electrical current through a wire is associated with energy loss. This is due to interaction of the electrons traveling through the wire with the atoms in the wire, whereby they give some of their energy to the atoms, heating up the wire. This phenomenon is called electrical resistance. Wherever we look around us, there is abundance of examples like this. However, superconductors are an exception, they do not exhibit electrical resistance [5–7]. Electrical current flows through superconductors without any loss, until it reaches a critical value. So are they not subject to the same physical laws as all the other phenomena around us? The answer to this can be found in the laws of quantum mechanics.

Quantum mechanics is a branch of physics applicable to atomic and subatomic particles. One of its main postulates is that the energy of these particles can have only certain discrete values [8–11]. The quantum state of a particle, or a system of particles, with the lowest allowed energy is called the ground state. The energy of a particle in the ground state can change only by an amount equal to the difference between the energy of higher lying states and the ground state. Therefore, there is a minimum energy required by the particle in the ground state in order to attain a state with the nearest higher energy level. On the other hand, it cannot transfer its energy to its surroundings, because it is already in the state with the lowest allowed energy level. These principles are vital for understanding the superconductivity and consequently for developing the idea of vortex pinning by nanoparticles in high-temperature superconductors.

The first model explaining superconductivity was the Bardeen–Cooper–Schriefer (BCS) model [12, 13]. It described superconductivity as a quantum state where electrons or holes form Cooper pairs. By pairing, they attain a lower energy level and they are in the ground state. The nearest available state is of a higher energy and there is an energy gap between these two states. When in the ground state, the charge carriers cannot exchange energy with their environment, as long as the energy of the environment is lower than this energy gap. Because of this, they can transfer electrical current through a wire without heating it up (i.e., there is no

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (207–218)

electrical resistance in this wire). However, if thermal energy of the environment is high enough to overcome the energy gap, the Cooper pairs will be broken up and the charge carriers will attain quantum state with a higher energy level. The temperature at which this occurs is called the critical temperature (T_c). BCS theory predicts that the energy gap is actually 3.52 times the thermal energy at T_c, which is close to the value for most of the known superconductors.

All this implies that the superconducting state is a state of lower energy than the resistive state. Therefore, this is a stable state of the materials exhibiting superconductivity and transition into the resistive state requires input of external energy.

Another characteristic property of superconductors is that they screen out the external magnetic field, so that the field $\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{M})$ is zero inside the superconducting volume [5–7]. Here, **B**, **H**, and **M** are vectors of the net magnetic field (also called magnetic induction), external field, and magnetization, respectively. $\mu_0 = 4\pi \times 10^{-7}$ T^2 m^3/J is the permeability of vacuum. This occurs because the magnetic field would otherwise break up the Cooper pairs and the total energy of the system would be higher if the field penetrated superconducting volume. Superconductors are divided into two groups, in regard to their response to external magnetic field. The type I superconductors screen out the external field completely, for small magnetic fields. This is accomplished by maintaining persistent supercurrent near the surface of the superconductor. These supercurrents produce internal field (or magnetization **M**), which cancels the field **H**. For higher fields, the thickness of the layer carrying the supercurrents gradually increases and superconductivity is finally broken down at a field called the critical field H_c.

The type II superconductors behave the same way as type I, up to the first critical field, H_{c1}. For higher fields, they are in a mixed state. Here, the field penetrates the superconductor in the form of thin filaments, called magnetic vortices [14] (Fig. 1). Each vortex carries magnetic flux of one flux quantum: $\Phi_0 = 2.067 \times 10^{-15}$ T m^2. The vortex consists of a core, which is in nonsuperconducting state. The diameter of the core is equal to the *coherence length*, ξ. Magnetic flux inside this core is screened out from the rest of the superconducting volume by persistent supercurrents, circulating the core. The value of the supercurrents decreases exponentially with the distance from the vortex center, with characteristic length scale equal to the London penetration depth, λ. This is mirrored in exponential decrease of **B** with distance from the core (Fig. 2). As external field increases, more magnetic vortices move into the superconductors from its surface. The overall screening around the whole of the superconductor, similar to the type I superconductors, still persists in the mixed state. When the external field reaches a value called the second critical field, H_{c2}, superconductivity is broken down.

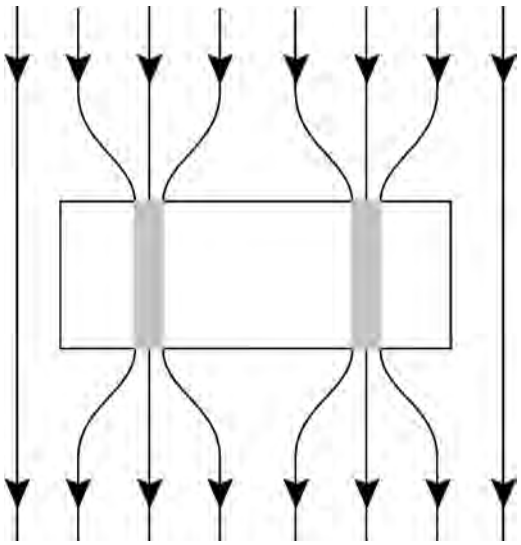


Figure 1. Schematic drawing of the field entering the type II superconductor through magnetic vortices (shaded).

3. VORTEX PINNING

In a perfect superconducting crystal, the vortices make a vortex lattice. The position of each of the vortices is defined by the repulsive force between them. When a current driven by external source is applied to this crystal, the vortices will be subjected to Lorentz force, because they carry a magnetic moment. Because of this, they will move through the crystal, exiting it on one side, whereas new vortices will be entering it on the other side, so that the total number of vortices in the crystal is always the same. The average Lorentz force per unit volume of the crystal is [15]

$$\mathbf{F}_L = \mathbf{J} \times \mathbf{B} \tag{1}$$

where **J** and **B** are spatially averaged transport current density and magnetic field inside the crystal, respectively.

Movement of the vortices is associated with a change of magnetic flux in and around the vortices [16]. This creates electromagnetic force that drives the supercurrents into the resistive core of the vortices. In effect, some of the Cooper pairs are broken down and the charge carriers interact with the crystal lattice, resulting in dissipation of their energy and electrical resistance. According to Bardeen and Stephen [16], about half of the dissipation occurs within the vortex core and half in the transition region just outside the core. This means that a perfect superconductor in the mixed state

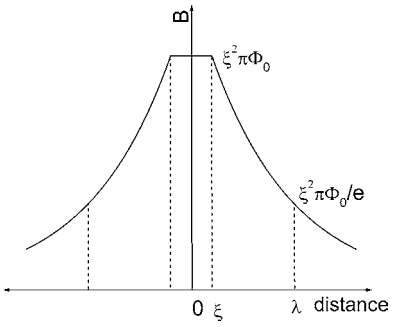


Figure 2. Spatial variation of **B** around a magnetic vortex. **B** is constant inside the nonsuperconducting core of radius of the coherence length ξ. It decreases approximately exponentially with distance outside the core, with the decay length equal to London penetration depth λ.

will exhibit electrical resistance regardless of the value of the transport current.

If, however, the superconducting crystal has defects in its structure, there will be local variations in the superconducting properties. This will result in energetically favorable sites for magnetic vortices, where they will be pinned down. Such sites are called the *pinning centers*. With the vortices pinned, there is again no electrical resistance in the superconductor. However, when applying too large a magnetic field or transport current, Lorentz force on the vortices will exceed the pinning force of the pinning centers. This will set vortices to motion again, and resistance will occur. The current at which the resistance just starts occurring is called the *critical current*, I_c. Additionally, if the thermal energy is higher than the pinning energy, vortices can also escape from the pinning centers. The thermal energy is proportional to temperature T,

$$E_{\text{th}} = k_B T \tag{2}$$

where k_B is the Boltzman constant. Because of this, I_c will become zero for certain combinations of magnetic field and temperature, for which vortices are just released from the pinning centers. They are called *irreversibility field* (H_{irr}) and *irreversibility temperature* (T_{irr}), respectively. Because $T_{\text{irr}} < T_c$ and $H_{\text{irr}} < H_{c2}$, electrical resistance in type II superconductors can occur even when they are still in the superconducting state. Values of H_{irr}, T_{irr}, and critical current density (J_c) are defined by the vortex pinning.

3.1. Mechanisms of Vortex Pinning

Magnetic vortices themselves are associated with variation of superconductivity in them. The core of the vortices is in nonsuperconducting state and carries a magnetic flux. Since the superconductivity is suppressed in the core, its volume is in a state with a higher energy than the superconducting matrix. The energy of the core per unit length is [17]

$$E_{\text{core}} = \frac{H_c^2}{2\mu_0}\pi\xi^2 \tag{3}$$

where $H_c = \Phi_0/(2\pi\xi\lambda\sqrt{2})$.

The core is surrounded by circulating supercurrents. The magnetic energy associated with the currents per unit length of the vortex was estimated as [17]

$$E_{\text{mag}} \approx \frac{H_c^2}{2\mu_0} 4\pi\xi^2 \ln\kappa \tag{4}$$

where $\kappa = \lambda/\xi$ is the Ginzburg–Landau parameter.

As the vortices move through the superconducting crystal with defects, they will encounter places in the crystal with suppressed superconductivity. The total energy of the whole superconducting system is then minimized. This is because vortices themselves are associated with suppression of superconductivity. Occupying a volume in the crystal where superconductivity is also suppressed, the net increase of energy due to the presence of the vortices is lower. Because of this, vortices will be pinned at such pinning centers. There are different types of basic interaction between the vortices and pinning centers.

3.1.1. Core Interaction

Pinning via *core interaction* occurs due to minimization of the core energy of the vortices when they encounter pinning centers [17, 18]. The vortex core energy saved when a vortex encounters a pinning center depends on the geometry of the pinning center and degree of suppression of the order parameter. Pinning force of a pinning center depends on the gradient of the energy variation within the pinning center. If the pinning center is much larger than the vortex core (i.e., ξ), it will be a weak pinning center. This is because the energy of the vortex will change very gradually as it passes the pinning center. On the other hand, if it is much smaller than ξ, it cannot contribute substantially to lowering the energy of the system and the energy gradient is again small, even though the variation occurs on a much smaller scale. Therefore, the most effective pinning centers are of the size of ξ.

3.1.2. Magnetic Interaction

When a vortex approaches a defect in the crystal, its circulating current and associated magnetic fields are redistributed because of the inhomogeneities around the defect. This is reflected in a change of magnetic energy of the vortex. Numerous reports have shown that this interaction can result in vortex pinning [19–21].

3.1.3. Interaction through Variation of Elastic Energy

Superconductivity is associated with small changes of volume ($\Delta V/V \approx 10^{-7}$) and elastic constants ($\Delta K/K \approx 10^{-4}$) of the crystal lattice [17] Because of this, the nonsuperconducting core of magnetic vortices will have different crystal lattice parameters and elastic constants. Consequently, they will behave like inclusions in the superconducting matrix, producing permanent elastic stress field around the vortices. This field will interact with crystal imperfections, which can result in pinning of the vortices [22–25].

3.2. High-Temperature Superconductors

High-temperature superconductors consist of metal-oxide layers, stacked atop each other along the crystalline *c*-axis [26, 27]. There are currently more than 100 known HTS systems. Critical temperature of different HTS spans from below 20 K for $Bi_2Sr_2CuO_6$ [28] or $(Nd_{0.925}Ce_{0.075})_2CuO_{4-\delta}$ [29] to 164 K for $HgBa_2Ca_2Cu_3O_{8+\delta}$ at high pressure [30]. Critical temperatures and coherence lengths of some HTSs are given in Table 1. For comparison, the highest values of T_c for classical superconductors are 23.2 K for Nb_3Ge [50, 51] and 18 K for Nb_3Sn [52].

A common feature of all HTSs is CuO_2 layers situated in the crystalline *ab*-plane (there are some exceptions, for example, $Ba_{1-a}K_aBiO_3$ with T_c of 32 K [53–56] and $YSr_2Ru_{1-b}Cu_bO_6$ with T_c of 60 K [57]). It is generally accepted that the primary sites of superconductivity in HTS crystal structure are these CuO_2 layers [27, 58–60], even though there are some other views, too [61]. While CuO_2 layers are a good conducting medium, the oxide layers between CuO_2 are poor conductors. This results in anisotropic superconducting properties of HTSs. The values

Table 1. Critical temperature (T_c) and in-plane (ξ_{ab}) and out-of-plane (ξ_c) coherence lengths of some high-temperature superconductors.

Superconductor	T_c (K)	Ref.	ξ_{ab} (nm)	ξ_c (nm)	Ref.
$YBa_2Cu_3O_7$	92	[31, 32]	1.2–1.8	0.2–0.3	[43–47]
$DyBa_2Cu_3O_7$	89	[33]			
$Bi_2Sr_2Ca_2Cu_3O_{10}$	110	[34]	0.63–2.9		[35, 36]
$Bi_2Sr_2CaCu_2O_8$	85	[37]	2–4	0.04	[38, 39, 48, 49]
$HgBa_2CuO_{4+\delta}$	97	[30]	2.1	1.2	[40]
$HgBa_2CaCu_2O_{6+\delta}$	128	[30]	1.7	0.4	[41]
$HgBa_2Ca_2Cu_3O_{8+\delta}$	135	[30]	1.5–2	0.19	[40–42]
$Tl_2Ba_2Ca_2Cu_3O_z$	125	[29]			

Note: All values given in the table are measured at atmospheric pressure.

of ξ and λ in the crystalline *ab*-layer are very different from the ones along the crystalline *c*-axis. For example, $\xi_a \approx \xi_b \approx$ 1.2–1.8 nm and $\xi_c \approx 0.2$–0.3 nm for $YBa_2Cu_3O_7$ superconductors (Table 1) [43–47, 62]. Technologically most important HTSs are currently $YBa_2Cu_3O_7$, $Bi_2Sr_2Ca_2Cu_3O_{10}$, and $Bi_2Sr_2CaCu_2O_8$, with anisotropy increasing in the same order.

The small coherence length of HTSs is a major cause for poor connectivity between the HTS crystals. This severely limits transport of the superconducting current through macroscopic conductors made of HTSs. The grain connectivity is still a major obstacle for wider introduction of HTSs in practical applications. Complexities of grain connectivity are beyond the scope of this chapter; detailed reviews can be found elsewhere [63]. The problem of grain connectivity was most successfully solved for $Bi_2Sr_2Ca_2Cu_3O_{10}$. It is produced in the form of kilometer-long wires [30], which can be used for electric cables, transformers, electrical motors, and generators. A weak point of this superconductor is strong suppression of its intracrystalline J_c by magnetic field. In contrast to this, J_c of $YBa_2Cu_3O_7$ crystals is much less sensitive to magnetic field; however, the grain connectivity is a serious problem. Current research efforts are aimed at developing thick films of $YBa_2Cu_3O_7$ with good grain connectivity over a large length. It is predicted that they will replace $Bi_2Sr_2Ca_2Cu_3O_{10}$-based wires in about 10 years.

3.2.1. Magnetic Vortices in High-Temperature Superconductors

The anisotropy in HTSs is reflected in a special structure of magnetic vortices in them. They are described as a pancake-shaped structures residing in CuO_2 layers, with weak connection between the pancake vortices in the neighboring CuO_2 layers (Fig. 3) [64, 65]. The higher the anisotropy of the superconductor, the weaker this connection. Because of a slight anisotropy in the *ab*-plane, the shape of the pancakes is elliptical. The vortex core is of the size of ξ_a and ξ_b in the direction of *a*- and *b*-axes, respectively. The screening currents circulate the core over the length scale of λ_a and λ_b in the direction of *b*- and *a*-axes, respectively.

The density of the vortices depends on magnetic field, because each of the vortices carries one magnetic flux quantum Φ_0. At elevated fields, the density of the vortices is high and they closely approach one another. Because of this, there is a strong interaction between the pancake vortices in each CuO_2 layer in high magnetic fields. In contrast to this, magnetic field strongly suppresses the interaction between the pancake vortices in neighboring layers [65].

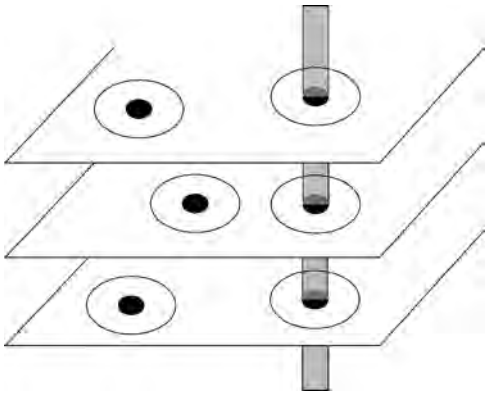


Figure 3. Schematic drawing of pancake vortices in high-temperature superconductors. The nonsuperconducting core (solid ellipse) of the size of coherence length ξ is surrounded by screening currents (open ellipse) that decay with the distance from the core, with a characteristic decay length λ. The diagram is not to scale, because $\xi \ll \lambda$ for high-temperature superconductors. The planes represent the CuO_2 layers. There is only weak interaction between the pancake vortices in neighboring layers and therefore there is almost no correlation of the distribution of pancake vortices in different layers. The shaded area is a columnar defect, which pins down the pancake vortices in different CuO_2 layers along a straight line.

This makes it very difficult to suppress the vortex motion. If there is no strong interaction between the vortices, pinning of one of them will not prevent the others from moving. Introducing the pinning center for each of the vortices is not a viable solution, because the large density of defects in the crystal would destroy the superconductivity in it. The only feasible way is to strengthen the interaction between the pancake vortices. By pinning one pancake vortex in an array of strongly interacting vortices, the unpinned vortices will be tied to the pinning center via this interaction.

3.2.2. Pinning Force in High-Temperature Superconductors

The magnetic energy of the vortices is larger than the core energy by a factor $4 \ln \kappa$ [Eqs. (3) and (4)]. The value of κ for high-temperature superconductors is of the order of 10^3–10^5, implying that the core energy of the vortices is much smaller than the magnetic energy. However, the pinning force depends on the gradient of the pinning energy. The core energy of the vortex changes on the length scale of ξ, and magnetic energy changes on the length scale of λ. Therefore, the gradients of the core energy and magnetic energy are proportional to $1/\xi$ and $1/\lambda$, respectively. This leads to the ratio of the core to magnetic pinning force as

$$\frac{F_{\text{core}}}{F_{\text{mag}}} \propto \frac{\kappa}{\ln \kappa} \tag{5}$$

Because for high-temperature superconductors $\kappa \gg 1$, F_{core} is always much larger than F_{mag}. Pinning force due to the variation of the elastic energy is negligible compared to F_{core} and F_{mag}.

In order to counteract the effect of magnetic field on J_c of high-temperature superconductors, the pinning force has to be larger than the Lorentz force. From this argument if follows that the only effective way of doing so is to improve the vortex pinning by introducing the pinning centers of the size of ξ. The value of ξ in high-temperature superconductors is less than ~10 nm [49–49, 66], which shows that nanosize pinning centers have to be employed in improvement of the critical currents in high-temperature superconductors.

4. NANOSIZE PINNING CENTERS

Nanosize pinning centers have two main advantages over the larger pinning centers. First, they provide a strong pinning force, because their size is comparable to ξ. Second, the volume with suppressed superconductivity is minimized, enabling introduction of a large number of pinning centers without significant degradation of superconductivity of the crystal. These pinning centers can be of spherical or columnar shape, each of them having different pinning properties. They are generally introduced into superconducting crystals by irradiation or doping by nanoparticles.

4.1. Irradiation Defects

Charged elementary particles or heavy ions interact with the atoms of superconducting crystals, transferring some of their kinetic energy to these atoms. If the transferred energy is larger than threshold energy, defects are formed in the crystal. Because each of the defects is formed by individual accelerated particles, the defect size is on the scale of 10 nm. Therefore, these defects are comparable in size to ξ, making them strong pinning centers. The type of defect depends strongly on the type and energy of irradiation.

4.1.1. Heavy Ion Irradiation

Accelerated heavy ions can interact with the target HTS crystal via elastic collisions with the target atoms or via inelastic interaction with the crystal lattice, resulting in ionization or excitation of electron states of the crystal. The type of the incurring damage will depend on the rate of energy loss, dE/dx, of the ions as they penetrate the crystal.

For low-energy ions, elastic collisions dominate, resulting in displacement of atoms from crystal lattice and pointlike defects. For high-energy ions, inelastic collisions dominate. Experiment shows that the inelastic collisions with high-energy ions can result in amorphous columnar tracks in HTSs [67]. The diameter of the tracks is about 10 nm and the length is tens of μm. An important parameter defining the formation of the columnar defects is the electronic energy loss, also called electronic stopping power, S_e. It is defined as the energy transfer into the electronic excitations of the target atoms per unit length along the ion path through the target crystal: $S_e = -dE_e/dx$. The value of S_e depends on the type of the target crystal, type, and energy of the ion and direction of the irradiation in regard to the crystal structure. It can be approximated as [68]

$$S_e = \frac{2h^3}{m_e \pi^2 e^2}\left(\frac{2E}{M}\right)^{1/2} Z^{2/3} \sum_i n_i Z_i^{1/3} \tag{6}$$

where Z and Z_i are respectively the atomic numbers of the incident ions and various target atoms, E is the energy of incident ions, n_i is the number of the target atoms of type i per unit volume, M and m_e are respectively the mass of the incident ion and electron, h is the Planck constant, and e is the base of natural logarithms.

Formation of amorphous columnar defects is described by the thermal spike model [69–72]. Because columnar defects are formed as a consequence of electronic excitation by the heavy ions, there has to be a mechanism of the transfer of this energy to the crystal lattice. The thermal spike model does not explain the nature of these interactions; however, it explains formation of the columnar defects in a thermodynamic approach. The energy transfer from the heavy ions to the electronic excitations occurs on a much smaller time scale than the subsequent transfer of this energy into the thermal energy of the crystal lattice. This heats up the crystal on a very localized scale around the atom targeted by a heavy ion. The heat is then diffused to the rest of the crystal on a larger time scale than the energy transfer to the crystal lattice, until the thermal equilibrium is reached.

Because of the transfer of the energy from the heavy ions to the electrons and to crystal lattice on a subsequently increasing time scale, a sudden localized increase of temperature occurs around the target atom of the crystal. For the value of S_e higher than a threshold value S_{e0}, the crystal lattice melts in much localized volume along the ion track. Because the diameter of the molten volume is very small (of the order of a nanometer), this heat is diffused through the crystal lattice quickly enough to freeze the molten volume without allowing recrystallization. Consequently, a track of amorphous material is formed in the crystal, which is not superconducting.

High-resolution electron microscopy analysis suggests that there are five ranges with different types of the damage, depending on S_e [67, 73–75]. Below S_{e0}, the damage occurs only due to the elastic collisions. For higher values of S_e, small spherical amorphous defects were observed along the ion track. These defects become elongated for yet higher S_e. For still higher S_e, discontinuous amorphous cylinders were observed along the ion track. For the largest values of S_e, continuous columnar amorphous cylinders were observed along the whole of the ion tracks. The values of S_e separating different ranges depend on the properties of the ions and target.

The thermal spike model gives a relationship between S_e and the track diameter. For the ion beam directed along the c-axis of the crystal, the amorphous track is of circular cross-section with diameter [68]

$$d = \left(\frac{4S_e}{\pi e C \Delta T_m}\right)^{1/2} \tag{7}$$

where C is the heat capacity per unit volume and ΔT_m is the difference between the melting and ambient temperature when the ion travels along the a- or b-axis of the crystal. If the ion travels along the a- or b-axis, the amorphous track is

of elliptical cross-section. The elliptical axes in the *ab*-plane and *c*-axis are respectively [68]

$$d_{ab} = d\left(\frac{D_{ab}}{D_c}\right)^{1/4} \tag{8}$$

$$d_c = d\left(\frac{D_c}{D_{ab}}\right)^{1/4} \tag{9}$$

Here, $D_{ab} = \sigma_{ab}/C$ and $D_c = \sigma_c/C$, with σ_{ab} and σ_c being the thermal conductivity in the crystalline *ab*-plane and along the *c*-axis, respectively. The value of d increases with the energy of the heavy ions via S_e and ΔT_m in Eq. (7).

4.1.2. Neutron and Proton Irradiation

Neutrons are electrically neutral particles and they do not have significant interaction with the electrons. They interact with the nuclei of the atoms in crystalline lattice. The maximum energy transferred to a crystal atom in elastic collision with a neutron is [76]

$$E_t = \frac{4m_n M}{(m_n + M)^2} E_n \tag{10}$$

where m_n and M are respectively the mass of neutron and target atom and E_n is the incident neutron energy. Neutron irradiation produces defects, with values of E_t ranging from the displacement threshold energy for different crystal lattices (~20 eV) to well above 100 keV.

For low values of E_t, point defects are produced. However, for roughly $E_t > 30$ keV, cascade defects are observed. When a high-energy neutron collides with an atom, E_t may be much higher than the threshold energy for the displacement of the atom. The high-energy displaced atom will cascade with other atoms in the crystal, displacing them, too. Cascade collisions will occur, eventually transferring the energy to the crystal lattice. This increases the temperature of the lattice beyond the melting point. A fast cooling follows, due to the small size of the molten volume, resulting in amorphous defects in the crystal. The size of these defects is 2–30 nm and it increases with the neutron energy [77–80]. The cascade defects are of irregular shape and are generally larger than the columnar defects obtained in heavy ion irradiation [81]. As opposed to the point defects, they are not annealed out at 300 °C [78].

Protons are charged particles and when accelerated to 800 MeV, they can also produce columnar tracks [82], similar to the heavy ions. Low energy protons (3 MeV) produce point defects, which have different pinning properties [83].

Neutron and proton irradiation can also be used for introduction of defects in inelastic interactions with nuclei in crystalline lattice [84, 85]. These interactions can provoke a nuclear reaction of the atoms in the crystalline lattice, or dopants. The products of the reaction are high-energy charged particles, which produce columnar tracks along their paths. The diameter of the tracks is typically less than 10 nm. These tracks are randomly orientated and they have proven to provide vortex pinning that is stronger than pinning on parallel columnar tracks [86, 87].

4.1.3. Electron Irradiation

High-energy electrons can be accelerated to energies of ~30 MeV [88]. Most of their energy is transferred to the superconducting crystal via electronic excitations of the target atoms. However, the energy of the electrons is too small to produce amorphous tracks of the type of heavy ion irradiation in this process. A small part of their energy is transferred to the crystal via elastic collisions with the atoms. It is estimated that about 40 eV is transferred to the target atom per elastic collision [88], which is of the order of the threshold energy needed to displace the target atoms from HTS crystal. Because of this, electron irradiation can produce only point defects [89] in the crystal, which do not contribute significantly to the vortex pinning [90, 91].

4.1.4. Columnar Defects

High-resolution electron microscopy examinations of heavy-ion irradiated HTS crystals showed a qualitative agreement with the thermal spike model [68]. $YBa_2Cu_3O_{6.3}$ and $Bi_2Sr_2Ca_2Cu_3O_{10}$ superconductors were irradiated with Au^{24+}, Ag^{21+}, Cu^{18+}, and Si^{13+} ions of 300, 276, 236, and 182 MeV, respectively. The energy loss was respectively 34.8, 23.8, 13.4, and 3.8 keV/nm. Only Au^{24+} ions produced continuous amorphous columns along the whole of their tracks. Each Au^{24+} ion produced a single columnar defect. However, only 50% of Ag^{21+} ions produced continuous columnar defects. Cu^{18+} ions produced only occasional columnar defect, whereas no columnar defects were observed for Si^{13+} ions. There was a threshold energy loss, below which no columnar defects were observed, for both types of superconductors. The cross-sectional area of the columns was circular for the ion beam directed along the crystalline *c*-axis, whereas it was of elliptical shape with the ion beam along either the *a*- or *b*-axis [68].

Continuous columnar defects introduced by heavy ions with energies higher than several hundreds of MeV are the best pinning centers in HTSs because they are about the size of ξ and they effectively tie together pancake vortices in different CuO_2 layers along a single straight line for each pinning center (Fig. 3). As opposed to the point defects, every pancake vortex along the line is pinned down. Thus pinned pancake vortices effectively behave as a vortex line. This is energetically more favorable than a stack of weakly coupled pancake vortices pinned randomly by point defects. The pancake vortices would meander inside the crystal from one point defect to the other, and the magnetic flux would occupy a larger volume in the crystal than a straight vortex line pinned by a columnar defect. Therefore, columnar defects provide the largest net pinning force. In addition, pancake vortices of such vortex lines cannot be individually excited away from the pinning site by thermal excitation. Consequently, pinning by columnar defects will be effective at high temperatures, which is not the case with the pointlike pinning centers. Improvement of the vortex pinning by columnar defects is the largest for HTSs with the highest anisotropy, because they exhibit the weakest interlayer coupling between the pancake vortices.

Improvement of J_c by introduction of columnar pinning centers has been observed for many HTSs [92–96]. Columnar defects were especially effective for highly anisotropic

HTSs. Irradiation of $Bi_2Sr_2CaCu_2O_8$ with 502 MeV ^{127}I ions resulted in the highest reported increase of J_c, up to 1000 times, as measured at 30 K and field of 2 T [97]. J_c for less anisotropic HTSs is usually improved 4–100 times, depending on the temperature, field, and HTS involved (Fig. 4) [83, 91].

To show that the improvement of J_c was indeed due to columnar defects, $YBa_2Cu_3O_7$ crystals were irradiated with 580 MeV $^{116}Sn^{30+}$ ions at an angle of 30° to the crystalline c-axis [98]. Field dependence of J_c was measured with the field oriented at 30° and −30° to the c-axis. A much higher value of J_c was obtained when the field was parallel to the ion beam, because vortices were pinned more effectively when aligned along the columnar tracks produced by the ion beam.

Improvement of J_c with the introduction of columnar defects is reflected in the increase of the values of H_{irr} and T_{irr}, which form the *irreversibility line* in the (H, T) diagram. Above the irreversibility line, J_c is negligibly small. Measurements of the irreversibility line for $YBa_2Cu_3O_7$ crystals showed that irradiation with 5.3 GeV Pb ions to a fluence of 7×10^9 ions/cm^2 produces a dramatic shift of the irreversibility line to higher fields and temperatures (Fig. 5) [99]. Experiments with 580 MeV Sn ions showed similar but less pronounced effects [98]. A strong increase of H_{irr} and T_{irr} was also obtained by irradiation of more anisotropic $Bi_2Sr_2CaCu_2O_8$ with 230 MeV Au^{14+} ions [95] and 6 GeV Pb ions [100].

Another important feature of pinning by columnar defects is that they give strong improvement of J_c at high temperatures, above $T = 0.5\ T_c$. While irradiation of $YBa_2Cu_3O_7$ crystals with 580 MeV Sn ions produced a drastic increase of J_c at 77 K (Fig. 4), point defects introduced by 3 MeV proton irradiation gave much weaker improvement of J_c at this temperature, especially at high fields [83].

Introduction of columnar defects is accompanied by an increase of the effective pinning energy of the crystal [101]. This in contrast to radiation induced point defects, which do not affect the pinning energy substantially [91, 102].

4.1.5. Splayed Defects

Theory shows that it is better to introduce columnar defects that are not parallel to each other, but splayed randomly in many directions [86, 103–107]. The most efficient way

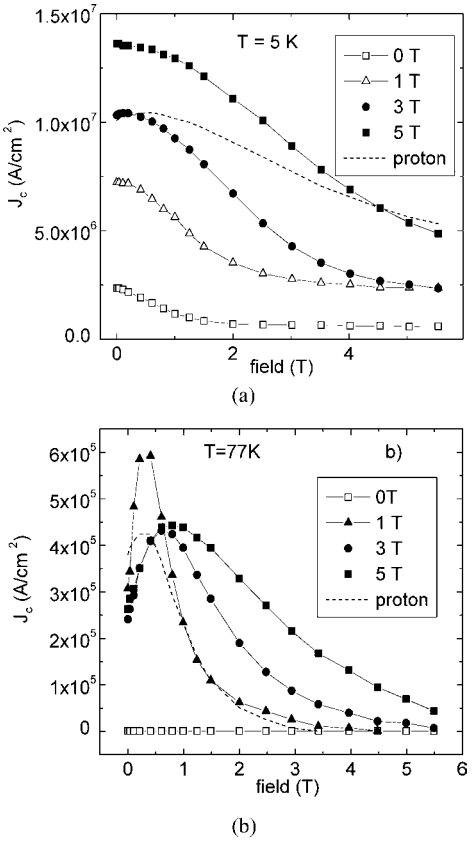


Figure 4. Critical current for $YBa_2Cu_3O_7$ crystals irradiated with 580 MeV Sn ions along their crystalline c-axis. The density of the defects produced is equivalent to trapping the magnetic vortices whose net magnetic field would correspond to the numbers indicated in the figures (each vortex carries magnetic flux $\Phi_0 = 2 \times 10^{-15}$ T m^2). Field dependence of J_c for the crystals irradiated with 3 MeV protons is shown for comparison. The temperature at which J_c was measured was (a) 5 and (b) 77 K. Reprinted with permission from [98], L. Civale et al., *Phys. Rev. Lett.* 67, 648 (1991). © 1998, American Physical Society.

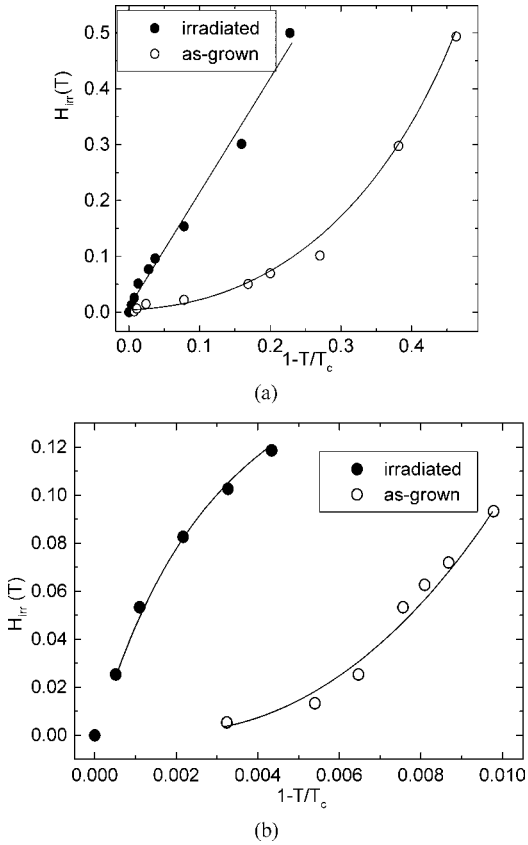


Figure 5. Irreversibility line for $YBa_2Cu_3O_7$ crystals before (open symbols) and after (solid symbols) irradiation with 5.3 GeV Pb ions. (a) The irreversibility line in a wide temperature range, down to 0.5 T_c. (b) Irreversibility line close to T_c. Reprinted with permission from [98], M. Konczykowski et al., *Phys. Rev. B* 44, 7167 (1991). © 1991, American Physical Society.

to introduce the splayed columnar defects is to instigate nuclear fission of small number of atoms in superconducting crystal. This can be achieved by doping an HTS with small amount of ^{235}U and then irradiating it with thermal neutrons [84, 108, 109]. ^{235}U is split into two fission products with energies of about 100 MeV each [110]. Because of the relatively low energy of the fission products, discontinuous columnar defects are obtained [111]. These defects are randomly splayed in the crystal. They produce an increase of J_c of 10–20 times for $Bi_2Sr_2Ca_2Cu_3O_{10}$ and shift of the irreversibility line to higher fields and temperatures.

Another way to introduce the splayed defects by fission fragments is irradiating $Bi_2Sr_2CaCu_2O_8$ that contains ^{209}Bi with 0.8 GeV protons [87, 112]. Irradiating ^{209}Bi with neutrons triggered nuclear fission. Typical fission products are 80 MeV Xe and 100 MeV Kr, which produce splayed discontinuous columnar defects of diameter up to 7 nm. A strong improvement of irreversibility line and J_c was observed, which was better than for the parallel columnar defects. The improvement was observed at all temperatures and fields. Similar results were obtained with neutron irradiation of lithium-doped $Bi_2Sr_2CaCu_2O_8$ [113]. There, neutron irradiation caused nuclear fission of Li atoms.

The effect of the splayed columnar defects on vortex pinning can be most effectively studied by rocking the HTS crystal upon high-energy heavy ion irradiation [114, 115]. As opposed to the defects introduced by fission products, the direction of the columnar defects can be easily controlled in this way. Increase of J_c by 14 times was obtained when irradiating $DyBa_2Cu_3O_7$ crystal at 45° and −45° to its surface, as compared to the parallel ion tracks with the ion beam perpendicular to the crystal surface [114]. The optimum splay angle for $YBa_2Cu_3O_7$ crystal was found to be ±5° with respect to the axis perpendicular to the *ab*-plane, with magnetic field oriented along this axis (i.e., at 0°) [116]. After irradiation with splay angle of ±5°, J_c increased respectively 10 and 10^4 times at temperatures of 5 and 77 K. With different splay angles, the increase was up to several times lower. However, splaying of the columnar defects seems to be much less effective for more anisotropic HTSs, like $Bi_2Sr_2CaCu_2O_8$ [100, 107]. This was attributed to the more pronounced pancake structure of magnetic vortices in highly anisotropic HTS, and increase of the diameter of columnar tracks by tilting the track away from the crystalline *c*-axis [100] [Eqs. (7)–(9)].

4.1.6. Point Defects

Point defects are introduced by low-energy particles (ions, protons, neutrons, electrons), which have low stopping power in HTS crystals. A common property of the pointlike pinning centers is that they generally improve J_c at temperatures lower than 0.5 T_c [88, 90, 91]. At higher temperatures, the improvement is insignificant. The pinning energy of these defects differs very little from the energy of the defects already in the crystal before the irradiation [91, 102]. Therefore, J_c increases because of the increased density of the pinning centers with the irradiation.

As opposed to the columnar pinning centers, the irreversibility line does not change with introduction of pointlike pinning centers. The irreversibility line for 3 MeV proton irradiated $YBa_2Cu_3O_7$ crystals was found to be independent of the proton fluence, even though J_c was improved by more than 10 times at 77 K [83]. On the other hand, irradiation of $Bi_2Sr_2CaCu_2O_8$ with 400 MeV oxygen ions resulted in only a slight shift of irreversibility line to higher values of fields and temperatures [117]. This is consistent with the dependence of the type of damage track on the energy loss [68]. Protons of 3 MeV produce point defects, which do not affect the irreversibility line. On the other hand, light oxygen ions produce a small number of columnar tracks, together with point defects, affecting the irreversibility line slightly.

High-energy neutrons and protons can produce cascade defects, which have vortex pinning properties somewhere between the point and columnar defects. Irradiation of $YBa_2Cu_3O_7$ crystals with 3 MeV protons resulted in defects, 30% of which were cascade defects [83]. J_c increased by an order of magnitude at both 5 and 77 K, whereas the pinning energy increased by only 25%. The irreversibility line, however, did not change. Irradiation of $YBa_2Cu_3O_7$ crystals with fast neutrons ($E_n > 0.1$ MeV) resulted in point and cascade defects. Value of J_c increased by an order of magnitude, at both low and high temperatures [77, 78, 80]. As opposed to the point defects, the cascade defects could not be annealed out at 300 °C [78]. The irreversibility line was also shifted to higher values of field and temperatures. Irradiation of $HgBa_2CuO_{4+\delta}$ [118] and $HgBa_2Ca_2Cu_3O_{8+\delta}$ [119–121] with fast neutrons ($E_n > 0.1$ MeV) also resulted in an increase of J_c, H_{irr}, and T_{irr} at low and high temperatures.

4.2. Doping with Nanoparticles

Doping with nonsuperconducting nanoparticles is another way to introduce pinning centers into HTSs. This way is more practical than the irradiation techniques, from the point of view of large-scale applications of HTSs. This is because irradiation techniques would be difficult to employ in industrial size applications. The nanoparticles have to satisfy certain criteria to be acceptable as pinning centers. Their presence should not affect the formation of the superconducting phase. They should not agglomerate in the superconductor, ensuring that the nonsuperconducting inclusions are of the size of ξ. The molten precursors for the superconductor should adhere to the surface of the nanoparticles, to ensure their incorporation into superconducting crystal. As with the irradiation-induced defects, the pinning properties depend on the shape of the particles.

4.2.1. Spherical Particles

Pinning centers formed by spherical particles incorporated into superconducting matrix have pinning properties similar to discontinuous defects introduced by irradiation. Chemically stable compounds are used, such as MgO [122, 123], ZrO_2 [124], and SiC [125]. They were incorporated into $Bi_2Sr_2Ca_2Cu_3O_{10}$, $Bi_2Sr_2CaCu_2O_8$, and $YBa_2Cu_3O_7$ superconductors, respectively. SiC nanoparticles were also used for doping of silver-sheathed $Bi_2Sr_2Ca_2Cu_3O_{10}$ tapes [126, 127]. Spherical particles of diameter about 20 nm were thoroughly mixed with precursor powder for production of HTSs. The superconductor was formed in thermal treatment, in which a liquid phase occurred before the formation of the superconductor. Because the liquid adhered to

the surface of the nanoparticles, they were incorporated into the superconducting crystal. Increase of J_c with this doping was usually less than 10 times at 5 K, whereas there was no significant increase at high temperatures [128]. The doping levels were generally kept below 5% by weight.

Doping of $YBa_2Cu_3O_7$ with SnO_2 and CeO_2 particles resulted in a moderate increase of J_c [129–131]. However, using nanosize particles did not have any advantage over the micrometer-size particles [129]. This was attributed to poor adherence of the molten phase of the superconductor to these particles.

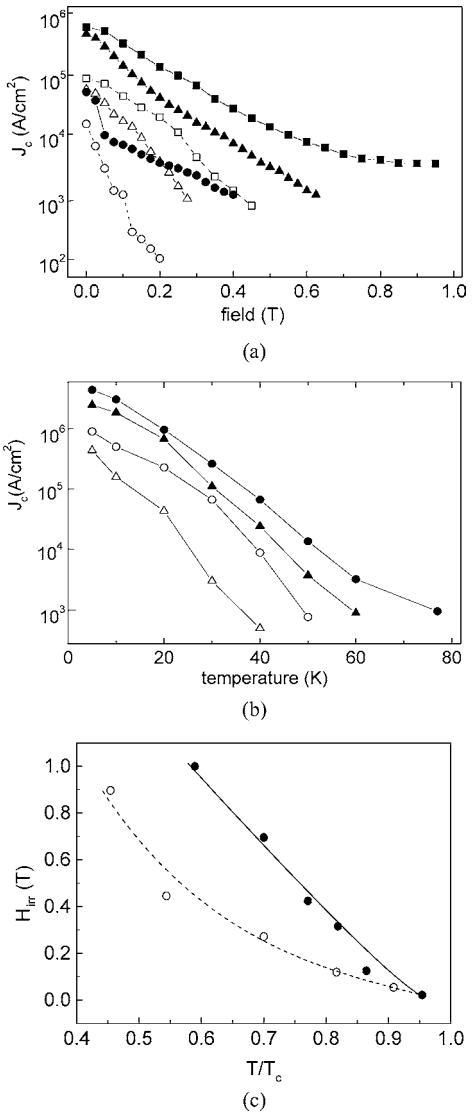


Figure 6. (a) Field dependence of critical current density for $Tl_2Ba_2Ca_2Cu_3O_z$ high-temperature superconductor doped with MgO nanorods. Measurements at 50, 60, and 90 K are represented by squares, triangles, and circles, respectively. Open and solid symbols represent J_c before and after introduction of MgO nanorods. (b) Temperature dependence of the critical current density for the same samples. The field was 0.5 and 0.8 T, represented by circles and triangles, respectively. (c) Irreversibility line for the same samples. Open and solid symbols represent the data before and after doping with MgO nanorods, respectively. Reprinted with permission from [133], P. Yang and C. M. Lieber, *Appl. Phys. Lett.* 70, 3158 (1977). © 1977, American Institute of Physics.

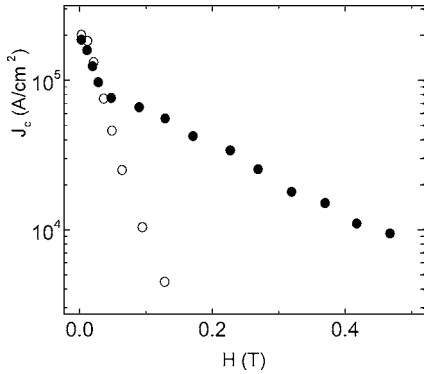


Figure 7. Field dependence of critical current density for $Bi_2Sr_2CaCu_2O_8$ crystals: pure (open symbols) and doped with nanosize CuO rods (solid symbols). Measurements were performed at 30 K. Reprinted with permission from [136], X. L. Wang et al., *J. Appl. Phys.* 81, 533 (1997). © 1997, American Institute of Physics.

4.2.2. Nanorods

Incorporating nonsuperconducting particles in the shape of nanorods should produce extended pinning centers resembling continuous columnar defects introduced by high-energy heavy ion irradiation. Nanorods of MgO were used for improving the vortex pinning in $Bi_2Sr_2Ca_2Cu_3O_{10}$, [132] $Bi_2Sr_2CaCu_2O_8$ [128], and $Tl_2Ba_2Ca_2Cu_3O_z$ [133] superconductors. The diameter of the nanorods was 5–50 nm and length was about 2 μm. They were orientated preferably along two directions in the crystal: parallel and perpendicular to the crystalline c-axis [128].

Doping increased the value of J_c about 10 times, at all temperatures (Fig. 6) [128, 132, 133]. The increase was more pronounced at higher fields and temperatures. The values of J_c depended approximately inversely on the diameter of the nanorods. The irreversibility line was shifted to higher fields and temperatures with the doping (Fig. 6). All this is consistent with the pinning by columnar defects. Further proof for that was doping of $Bi_2Sr_2CaCu_2O_8$ with MgO nanoparticles of spherical shape [128]. This resulted in an increase of J_c only at temperatures lower than 25 K, consistent with point-defect pinning.

$Bi_2Sr_2CaCu_2O_8$ was also doped with *carbon nanotubes* [134, 135]. The carbon nanotubes consisted of one or more cylindrical shells of graphite. Their outer and inner diameters were respectively 2–20 and 1–3 nm. Their length was several micrometers. Embedding the nanotubes into $Bi_2Sr_2CaCu_2O_8$ gave increased vortex pinning at temperatures lower than 53 K. This increase of pinning was ascribed to the nanotubes, which are of size and shape comparable to the irradiation induced columnar defects.

Needle-shaped particles of CuO were also incorporated into $Bi_2Sr_2CaCu_2O_8$ [136]. The CuO particles were of diameter of about 50 nm and length 3 μm. J_c increased by more than an order of magnitude at 30 K and $H > 0.2$ T (Fig. 7). Such large improvement of J_c was attributed to the small size of the particles and even distribution within the crystal.

5. CONCLUSION

Strong improvement of vortex pinning in HTSs can only be achieved by introduction of pinning centers smaller than 10 nm. For pinning centers of spherical shape, J_c improves substantially only at temperatures lower than about 0.5 T_c and the irreversibility line does not change significantly. Pinning centers in the form of continuous columns improve J_c at all temperatures and significantly shift the irreversibility line to higher fields and temperatures. These pinning centers are especially effective if they are oriented randomly in the superconducting crystal.

Nanosize pinning centers have been introduced into HTSs by irradiation and doping with nanoparticles. The strongest improvement of the vortex pinning was achieved by swift heavy ion irradiation, upon which continuous columnar defects were formed in HTS crystals. The most successful doping technique was utilization of MgO nanorods.

GLOSSARY

Coherence length Characteristic length scale over which the charge carriers within Cooper pairs can still interact with each other.

Cooper pairs Charge carriers paired together by exchange of a phonon. They are basic blocks of superconductivity.

Critical current Maximum current through a superconductor that still does not produce dissipation of energy.

Flux quantum The lowest possible value of magnetic flux. Any change of magnetic flux occurs stepwise, in multiples of the flux quantum.

Irreversibility line Field and temperature above which the critical current is zero.

ACKNOWLEDGMENT

Many thanks are due to my wife, Greer Taylor, for her help with preparation of the figures.

REFERENCES

1. J. G. Bednorz and K. A. Müller, *Z. Phys. B* 64, 189 (1986).
2. M. K. Wu, J. R. Ashburn, C. J. Torng, P. H. Hor, R. L. Meng, L. Gao, Z. J. Huang, Y. Q. Wang, and C. W. Chu, *Phys. Rev. Lett.* 58, 908 (1987).
3. R. B. VanDover, R. Cava, B. Batlogg, and E. Rietman, *Phys. Rev. B* 35, 5337 (1987).
4. J. M. Tarascon, L. H. Greene, W. R. McKinnon, G. W. Hull, and T. H. Geballe, *Science* 235, 1373 (1987).
5. A. C. Rose-Innes and E. H. Rhoderick, "Introduction to Superconductivity." Pergamon, New York, 1994.
6. C. G. Kuper, "An Introduction to the Theory of Superconductivity." Clarendon Press, Oxford, 1968.
7. V. L. Ginzburg and E. A. Andryushin, "Superconductivity." World Scientific, Singapore/New Jersey/London/Hong Kong, 1994.
8. L. J. Schiff, "Quantum Mechanics." McGraw–Hill, New York/St. Louis/San Francisco/Toronto/London/Sydney, 1968.
9. A. Messiah, "Quantum Mechanics," Vol. I (translated by G. M. Temmer). North-Holland, Amsterdam, 1965.
10. W. Pauli, "General Principles of Quantum Mechanics" (translated by P. Achtan and K. VenKatesan). Springer-Verlag, Berlin, 1980.
11. D. Home, "Conceptual Foundations of Quantum Physics: An Overview from Modern Perspectives." Plenum Press, New York, 1997.
12. J. Bardeen, L. N. Cooper, and J. R. Schrieffer, *Phys. Rev.* 108, 1175 (1957).
13. J. H. Schrieffer, "Theory of Superconductivity," pp. 10–23. Benjamin, New York, 1964.
14. A. A. Abrikosov, *Soviet Phys. JETP* 5, 117 (1957).
15. J. Friedel, P. G. de Gennes, and J. Matricon, *Appl. Phys. Lett.* 2, 119 (1963).
16. J. Bardeen and M. J. Stephen, *Phys. Rev.* 140, A1197 (1965).
17. S. Sengupta and D. Shi, in "High Temperature Superconducting Materials Science and Engineering: New Concepts and Technology" (D. Shi, Ed.), p. 131. Pergamon, Elmsford, NY, 1995.
18. L. D. Cooley and A. M. Grishin, *Phys. Rev. Lett.* 74, 2788 (1995).
19. G. S. Mktrchyan and V. V. Shmidt, *Sov. Phys. JETP* 34, 195 (1972).
20. W. E. Timms and D. G. Walmsley, *Phys. Status Solidi B* 71, 741 (1975).
21. A. Buzdin and D. Feinberg, *Physica C* 256, 303 (1996).
22. R. L. Fleischer, *Phys. Rev. Lett.* 3, 111 (1962).
23. W. W. Webb, *Phys. Rev. Lett.* 11, 191 (1963).
24. L. E. Toth and I. P. Pratt, *Appl. Phys. Lett.* 4, 75 (1964).
25. E. J. Kramer and C. L. Bauer, *Philos. Mag.* 15, 1189 (1967).
26. C. P. Poole Jr., T. Data, and H. A. Farach, "Copper Oxide Superconductors." Wiley, New York/Chichester/Brisbane/Toronto/Singapore, 1988.
27. F. Izumi and E. Takayama-Muromachi, in "High Temperature Superconducting Materials Science and Engineering: New Concepts and Technology" (D. Shi, Ed.), p. 81. Pergamon, Elmsford, NY, 1995.
28. C. Michel, M. Hervieu, M. M. Borel, A. Grandin, F. Deslandes, J. Provost, and B. Raveau, *Z. Phys. B* 68, 421 (1987).
29. A. Junod, in "Physical Properties of High-Temperature Superconductors II" (D. M. Ginsberg, Ed.). World Scientific, Singapore, 1990.
30. L. Gao, Y. Y. Xue, F. Chen, Q. Xiong, R. L. Meng, D. Ramirez, C. W. Chu, J. H. Eggert, and H. K. Mao, *Phys. Rev. B* 50, 4260 (1994).
31. S. J. Collocoot, R. Driver, and E. R. Vance, *Phys. Rev. B* 41, 6329 (1990).
32. R. P. Gupta and M. Gupta, *Phys. Rev. B* 51, 11760 (1995).
33. W. Carrilo-Cabrera and W. Gopel, *Physica C* 161, 373 (1989).
34. S. Chu and M. E. McHenry, *J. Mater. Res.* 13, 589 (1998).
35. W. M. Chen, F. Wang, S. S. Jiang, H. K. Liu, and S. X. Dou, *J. Supercond.* 14, 465 (2001).
36. I. Matsubara, H. Tanigawa, T. Ogura, H. Yamashita, M. Kinoshita, and T. Kawai, *Phys. Rev. B* 45, 7414 (1992).
37. K. Yvon and M. Francols, *Z. Phys. B* 76, 413 (1989).
38. B. vom Hedt, W. Lisseck, K. Westerholt, and H. Bach, *Phys. Rev. B* 49, 9898 (1994).
39. M. J. Naughton, R. C. Yu, P. K. Davies, J. E. Fischer, R. V. Chamberlin, Z. Z. Wang, T. W. Jing, N. P. Ong, and P. M. Chaikin, *Phys. Rev. B* 38, 9280 (1988).
40. R. Puzniak, R. Usami, and H. Yamauchi, *Phys. Rev. B* 53, 86 (1996).
41. R. Puzniak, R. Usami, K. Isawa, and H. Yamauchi, *Phys. Rev. B* 52, 3756 (1995).
42. Y. C. Kim, J. R. Thompson, J. G. Ossandon, D. K. Christen, and M. Paranthaman, *Phys. Rev. B* 51, 11767 (1995).
43. R. J. Cava, B. Batlogg, R. B. van Dover, D. W. Murphy, S. Sunshine, T. Siegrist, J. P. Remeika, E. A. Rietman, S. Zahurak, and G. P. Espinosa, *Phys. Rev. Lett.* 58, 1676 (1987).
44. T. K. Worthington, W. J. Gallagher, and T. R. Dinger, *Phys. Rev. Lett.* 59, 1160 (1987).

45. D. E. Farrell, C. M. Williams, S. A. Wolf, N. P. Bansal, and V. G. Kogan, *Phys. Rev. Lett.* 62, 2805 (1988).
46. L. Krusin-Elbaum, R. L. Greene, F. Holtzberg, A. P. Malozemoff, and Y. Yeshurun, *Phys. Rev. Lett.* 62, 217 (1989).
47. U. Welp, W. K. Kwok, G. W. Crabtree, K. G. Vandervoort, and J. Z. Liu, *Phys. Rev. Lett.* 62, 1908 (1989).
48. T. T. M. Palstra, B. Batlogg, L. F. Schneemeyer, R. B. van Dover, and J. V. Waszczak, *Phys. Rev. B* 38, 5102 (1988).
49. J. C. Martínez, S. H. Brongersma, A. Koshelev, B. Ivlev, P. H. Kes, R. P. Griessen, D. G. de Groot, Z. Tarnavski, and A. A. Menovsky, *Phys. Rev. Lett.* 69, 2276 (1992).
50. J. R. Gavaler, *Appl. Phys. Lett.* 23, 480 (1973).
51. L. R. Testardi, J. H. Wernick, and W. A. Royer, *Solid State Commun.* 15, 1 (1974).
52. H. K. Liu, M. Ionescu, and Y. C. Guo, in "Handbook of Electronic and Photonic Materials and Devices" (H. S. Nalwa, Ed.), Vol. 3, p. 71. Academic Press, San Diego/San Francisco/New York/Boston/London/Sydney/Tokyo, 2001.
53. L. F. Mattheiss, E. M. Gyorgy, and D. W. Johnson, *Phys. Rev. B* 37, 3745 (1988).
54. R. J. Cava, B. Batlogg, J. J. Krajewski, R. Farrow, L. W. Rupp Jr., A. E. White, K. Short, W. F. Peck, and T. Kometani, *Nature* 3332, 914 (1988).
55. R. G. Goodrich, C. Grienier, D. Hall, A. Lacerda, E. G. Haanappel, D. Rickel, T. Northington, R. Schwarz, F. M. Mueller, D. D. Koeling, J. Vuillemin, L. Van Bockstal, M. L. Norton, and D. H. Lowdness, *J. Phys. Chem. Solids* 54, 1251 (1993).
56. M. Licheron and F. Gervais, *J. Alloys Compd.* 195, 77 (1993).
57. M. K. Wu, D. Y. Chen, F. Z. Chien, S. R. Sheen, D. C. Ling, C. Y. Tai, G. Y. Tseng, D. H. Chen, and F. C. Zhang, *Z. Phys. B* 102, 307 (1997).
58. C. C. Torardi, M. A. Subramanian, J. C. Calabrese, J. Gopalakrishnan, E. M. McCarron, K. J. Morrissey, T. R. Askew, R. B. Flippen, U. Chowdhury, and A. W. Sleight, *Science* 240, 631 (1988).
59. J. T. Markert, Y. Dalichaouch, and M. B. Maple, in "Physical Properties of High Temperature Superconductors I" (D. M. Ginsberg, Ed.), p. 265. World Scientific, Singapore/New Jersey/London/Hong Kong, 1989.
60. T. M. Rice, *Physica C* 282–287, xix (1997).
61. A number of experiments would suggest that the prime sites of superconductivity are atomic layers containing hypocharged oxygen; see H. A. Blackstead and J. D. Dow, *J. Appl. Phys.* 83, 1540 (1998) and references therein. However, this model still needs further experimental verification in order to gain wider acceptance.
62. D. R. Harsham, L. F. Schneemeyer, J. V. Waszczak, G. Aepli, R. J. Cava, B. Batlogg, L. W. Rupp, E. J. Ansaldo, and D. L. Williams, *Phys. Rev. B* 39, 851 (1991).
63. J. Horvat, in "Handbook of Electronic and Photonic Materials and Devices" (H. S. Nalwa, Ed.), Vol. 3, p. 129. Academic Press, San Diego/San Francisco/New York/Boston/London/Sydney/Tokyo, 2001.
64. J. R. Clem, *Physica C* 162–164, 1137 (1989).
65. J. R. Clem, *Phys. Rev. B* 43, 7837 (1991).
66. N. Chikumoto, M. Konczykowski, N. Motohira, and K. Kishio, *Physica C* 199, 32 (1992).
67. J. Provost, Ch. Simon, M. Hervieu, D. Groult, V. Hardy, F. Studer, and M. Toulemonde, *Mater. Res. Soc. Bull.* 20, 22 (1995).
68. Yimei Zhu, Z. X. Cai, R. C. Budhani, M. Suenaga, and D. O. Welch, *Phys. Rev. B* 48, 6436 (1993).
69. F. Seitz and J. F. Koehler, in "Solid State Physics: Advances in Research and Applications" (F. Seitz and D. Turnbull, Eds.), Vol. 2, p. 305. Academic, New York, 1956.
70. G. Szenes, *Phys. Rev. B* 51, 8026 (1995).
71. L. T. Chadderton and I. M. Torrens, "Fission Damage in Crystals." Methuen, London, 1969.
72. M. Toulemonde, C. Dufour, and E. Paumier, *Phys. Rev. B* 46, 14362 (1992).
73. C. Houpert, F. Studer, D. Groult, and M. Toulemonde, *Nucl. Instrum. Methods B* 39, 720 (1989).
74. M. Toulemonde, G. Fuchs, N. Nguyen, F. Studer, and D. Groult, *Phys. Rev. B* 35, 6560 (1987).
75. M. Toulemonde, N. Enault, J.-Y. Fan, and F. Studer, *J. Appl. Phys.* 68, 1545 (1990).
76. Y. Kazumata, S. Okayasu, and H. Kumakura, in "Bismuth-Based High-Temperature Superconductors" (H. Maeda and K. Togano, Eds.), p. 174. Dekker, New York/Basel/Hong Kong, 1996.
77. F. M. Sauerzopf, H. P. Wiesinger, W. Kritscha, H. W. Weber, G. W. Crabtree, and J. Z. Liu, *Phys. Rev. B* 43, 3091 (1991).
78. B. M. Vlcek, M. C. Frischherz, S. Fleshler, U. Welp, J. Z. Liu, J. Downey, K. G. Vandervoort, G. W. Crabtree, M. A. Kirk, J. Giapintzakis, and J. Farmer, *Phys. Rev. B* 46, 6441 (1992).
79. M. A. Kirk, *Cryogenics* 33, 235 (1993).
80. H. W. Weber, *Supercond. Sci. Technol.* 5, S19 (1992).
81. M. A. Kirk, M. C. Frischherz, J. Z. Liu, L. R. Greenwood, and H. W. Weber, *Philos. Mag. Lett.* 62, 41 (1990).
82. H. Safar, J. H. Cho, S. Fleshler, M. P. Maley, J. O. Willis, J. Y. Coutler, J. L. Ullmann, P. W. Liowski, G. N. Riley, M. W. Rupich, J. R. Thompson, and L. Krusin-Elbaum, *Appl. Phys. Lett.* 67, 130 (1995).
83. L. Civale, A. D. Marwick, M. W. McElfresh, T. K. Worthington, F. H. Holtzberg, J. R. Thompson, and M. A. Kirk, *Phys. Rev. Lett.* 65, 1164 (1990).
84. H. R. Hart, Jr., F. E. Luborsky, R. H. Arendt, R. L. Fleishcher, J. E. Tkaczyk, and D. A. Orsini, *IEEE Trans. Magn.* 27, 1375 (1991).
85. L. Krusin-Elbaum, J. R. Thompson, R. Wheeler, A. D. Marwick, C. Li, S. Patel, D. T. Shaw, P. Lisowski, and J. Ullmann, *Appl. Phys. Lett.* 64, 3331 (1994).
86. T. Hwa, P. Le Doussal, D. R. Nelson, and V. M. Vinokur, *Phys. Rev. Lett.* 71, 3545 (1993).
87. L. Krusin-Elbaum, J. R. Thompson, R. Wheeler, A. D. Marwick, C. Li, S. Patel, D. T. Shaw, P. Lisowski, and J. Ullmann, *Appl. Phys. Lett.* 64, 3331 (1994).
88. T. Terai, T. Masegi, K. Kusagaya, Y. Takahashi, K. Kishio, N. Motohira, and K. Nakatani, *Physica C* 185–189, 2383 (1991).
89. J. Giapintzakis, W. C. Lee, J. P. Rice, D. M. Ginsberg, I. M. Robertson, R. Wheeler, M. A. Kirk, and M.-O. Ruault, *Phys. Rev. B* 45, 10677 (1992).
90. K. Shiraishi, Y. Kazumata, and T. Kato, *Jpn. J. Appl. Phys.* 30, L578 (1991).
91. H. Kumakura, H. Kitaguchi, K. Togano, H. Maeda, J. Shimoyama, S. Okayasu, and Y. Kazumata, *J. Appl. Phys.* 74, 451 (1993).
92. L. Civale, A. D. Marwick, R. Wheeler, M. A. Kirk, W. L. Carter, G. N. Riley, and A. P. Malozemoff, *Physica C* 208, 137 (1993).
93. J. R. Thompson, Y. R. Sun, H. R. Kerchner, D. K. Christen, B. C. Sales, B. C. Chakoumakos, A. D. Marwick, L. Civale, and J. O. Thompson, *Appl. Phys. Lett.* 60, 2306 (1992).
94. B. Chenevier, S. Ikeda, H. Kumakura, K. Togano, S. Okayasu, and Y. Kazumata, *Jpn. J. Appl. Phys.* 31, L777 (1991).
95. Y. Kazumata, X. Gao, H. Kumakura, and K. Togano, *Surface Coat. Technol.* 84, 348 (1996).
96. S. Khatua, P. K. Mishra, R. Kumar, V. C. Sahni, and R. Pinto, *Supercond. Sci. Technol.* 15, 324 (2002).
97. P. Kummeth, H. W. Neumüller, G. Ries, M. Kraus, S. Klaumünzer, and G. Saemann-Ischenko, *J. Alloys Compounds* 195, 403 (1993).
98. L. Civale, A. D. Marwick, T. K. Worthington, M. A. Kirk, J. R. Thompson, L. Krusin-Elbaum, Y. Sun, J. R. Clem, and F. Holtzberg, *Phys. Rev. Lett.* 67, 648 (1991).
99. M. Konczykowski, F. Rullier-Albenque, E. R. Yacoby, A. Shaulov, Y. Yeshurun, and P. Leyaj, *Phys. Rev. B* 44, 7167 (1991).
100. J. Provost, V. Hardy, Ch. Simon, M. Hervieu, D. Groult, and S. Hebert, *Int. J. Inorganic Mater.* 2, 635 (2000).

101. H. R. Kerchner, J. R. Thompson, Y. R. Suns, D. K. Christen, J. O. Thomson, B. C. Sales, B. Chakoumakos, L. Civale, and A. D. Marwick, *Physica B* 194, 1903 (1994).
102. H. W. Neumüller, G. Ries, W. Schmidt, W. Gerhauser, and S. Klaumünzer, *J. Less Common Metals* 164–165, 1351 (1990).
103. D. R. Nelson and S. Seung, *Phys. Rev. B* 39, 9153 (1989).
104. E. H. Brandt, *Phys. Rev. Lett.* 69, 1105 (1992).
105. E. H. Brandt, *Europhys. Lett.* 18, 635 (1992).
106. D. R. Nelson and V. M. Vinokur, *Phys. Rev. Lett.* 68, 2389 (1993).
107. Th. Schuster, H. Kuhn, M. V. Indenbom, M. Leghissa, M. Kraus, and M. Konczykowski, *Phys. Rev. B* 51, 16358 (1995).
108. E. Babić, I. Kušević, D. Marinaro, S. X. Dou, J. Boldeman, and R. Weinstein, *Solid State Commun.* 118, 607 (2001).
109. D. Marinaro, S. X. Dou, J. Horvat, J. Boldeman, R. Weinstein, and R. Sawh, *IEEE Appl. Supercond.* 11, 3764 (2001).
110. R. Weinstein, R. Sawh, Y. Ren, M. Eisterer, and H. W. Weber, *Supercond. Sci. Technol.* 11, 959 (1998).
111. A. Gandini, R. Weinstein, Y. R. Ren, R. P. Sawh, D. Parks, Y. C. Guo, B. Zeimetz, S. X. Dou, S. Tönies, C. Klein, and H. W. Weber, *Physica C* 341–348, 1453 (2000).
112. J. G. Ossandon and J. R. Thompson, *Superlatt. Microstruct.* 23, 543 (1998).
113. J. Schwartz and S. Wu, *J. Appl. Phys.* 73, 1343 (1993).
114. Th. Schuster, H. Kuhn, M. V. Indenbom, G. Kreiselmeyer, M. Leghissa, and S. Klaumünzer, *Phys. Rev. B* 53, 2257 (1996).
115. H. Kuhn and S. Klaumünzer, *Nucl. Instrum. Methods B* 146, 565 (1998).
116. L. Krusin-Elbaum, A. D. Marwick, R. Wheeler, C. Field, V. M. Vinokur, G. K. Leaf, and M. Palumbo, *Phys. Rev. Lett.* 76, 2563 (1996).
117. A. Gupta, P. Esquinazi, H. F. Braun, H.-W. Neumüller, G. Ries, W. Schmidt, and W. Gerhauser, *J. Phys. C* 170, 95 (1990).
118. J. Schwartz, S. Nakamae, G. W. Raban Jr., J. K. Heuer, S. Wu, J. L. Wagner, and D. G. Hinks, *Phys. Rev. B* 48, 9932 (1993).
119. A. Wisniewski, R. Puzniak, J. Karpinski, J. Hofer, R. Szymczak, M. Baran, F. M. Sauerzopf, R. Molinski, E. M. Kopnin, and J. R. Thompson, *Phys. Rev. B* 61, 791 (2000).
120. J. Krelaus, M. Reder, J. Hoffmann, and H. C. Freyhardt, *Physica C* 314, 81 (1999).
121. E. Altshuler, C. W. Chu, M. T. D. Orlando, A. Sin, A. J. Batista-Leyva, V. Buntar, and H. W. Weber, *Physica C* 224 (2002).
122. X. Wan, Y. Sun, W. Song, K. Wang, L. Jiang, and J. Du, *Physica C* 307, 46 (1998).
123. H. Sasakura, M. Osuke, and D. Ito, *IEEE Trans. Appl. Supercond.* 9, 2332 (1999).
124. Z. Y. Jia, H. Tang, Z. Q. Yang, Y. T. Xing, Y. Z. Wang, and G. W. Qiao, *Physica C* 337, 130 (2000).
125. Z. Q. Yang, X. D. Su, C. Zhang, G. W. Qiao, and W. Han, *Phys. Status Solidi A* 167, 165 (1998).
126. Z. Q. Yang, X. D. Su, G. W. Qiao, Y. C. Guo, S. X. Dou, and F. R. de Boer, *Physica C* 325, 136 (1999).
127. Y. C. Guo, Y. Tanaka, T. Kuroda, S. X. Dou, and Z. Q. Yang, *Physica C* 311, 65 (1999).
128. P. Yang and C. M. Lieber, *Science* 273, 1836 (1996).
129. S. Marinel and G. Desgardin, *J. European Ceramic Soc.* 21, 1919 (2001).
130. K. Osamura, N. Matsukura, Y. Kusumoto, S. Ochiai, B. Ni, and T. Matsushita, *Jpn. J. Appl. Phys.* 29, L1621 (1990).
131. S. Marinel, I. Monot, J. Provost, and G. Desgardin, *Supercond. Sci. Technol.* 11, 563 (1998).
132. P. D. Yang and C. M. Lieber, *J. Mater. Res.* 12, 2981 (1997).
133. P. Yang and C. M. Lieber, *Appl. Phys. Lett.* 70, 3158 (1997).
134. Sun-Li Huang, M. R. Koblischka, K. Fossheim, T. W. Ebbesen, and T. H. Johansen, *Physica C* 311, 172 (1999).
135. K. Fossheim, E. D. Tuset, T. W. Ebbesen, M. M. J. Treacy, and J. Schwartz, *Physica C* 248, 195 (1995).
136. X. L. Wang, J. Horvat, H. K. Liu, and S. X. Dou, *J. Appl. Phys.* 81, 533 (1997).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanopipes in Transition Metal Nitrides

Daniel Gall

Rensselaer Polytechnic Institute, Troy, New York, USA

CONTENTS

1. INTRODUCTION

Transition metal (TM) nitrides are well known for their remarkable physical properties including high hardness and mechanical strength, chemical inertness, and electrical resistivities that vary from metallic to semiconducting. As a result, they are widely studied and have become technologically important for applications such as hard wear resistant coatings on cutting tools, as diffusion barriers in microelectronic devices, and as corrosion and abrasion resistant layers on optical components. Transition metals from the left side of the periodic table, including Sc, Ti, V, Cr, Y, Zr, Nb, Hf, and Ta, form nitrides with a B1–NaCl structure [1, 2]. The excellent mechanical properties of these materials are due to strong covalent–ionic bonds between the TM and N ions resulting from the fully occupied N 2p bands [2, 3].

The chemical bonding in NaCl structure transition metal nitrides occurs along perpendicular $\langle 100 \rangle$ directions. This strong bonding directionality results in dramatic changes in surface atom mobilities as a function of crystalline facet orientation, leading to a range of highly anisotropic effects including strong preferred texture development in polycrystalline thin films, anisotropic microstructures, and orientation-dependent properties including hardness, wear strength, adhesion, chemical inertness, and diffusion.

Recent advances in the growth of single crystalline layers of NaCl structure transition metal nitrides have shown that these highly anisotropic effects cause another remarkable feature: Epitaxial layers exhibit, due to atomic shadowing from kinetically roughened periodic surface mound structures, arrays of self-organized rectangular 1-nm-wide nanopipes. These nanopipes are open "holes" that extend through the entire layer thickness along the [001] growth direction. Their high degree of self-organization, due to kinetic faceting, leads to extremely high aspect ratios on the order of 1000. This, in turn, offers tremendous potential for technological applications in a wide range of fields including nanoelectronics, single molecule chemistry, sensors, and integrated low-dimensionality devices.

This chapter summarizes current understanding for the growth of such nanopipes, the underlying physical properties of their formation, and methods to control their shape, size, and separation. The first section reviews recent findings describing the atomistic processes that govern the growth of transition metal nitrides, including the effects of anisotropic surface diffusion, atomic N flux, adatom potential energies, and ion irradiation. This is followed by a discussion on the key processes leading to nanopipe formation, that is, kinetic roughening, surface morphological evolution, and atomic shadowing. The last section describes the actual growth on nanopipes, strongly utilizing the concepts described in the preceding sections, and provides insight into nanopipe manipulation.

2. ATOMISTIC PROCESSES IN THE GROWTH OF TRANSITION METAL NITRIDE LAYERS

Transition metal nitride layers are typically grown by reactive sputter deposition, where the metal atoms, emanating from the sputter source, impinge on the substrate from the gas phase while the nitrogen supply is provided by a N_2 atmosphere. The microstructure of the deposited layers defined by their density, texture, strain, grain sizes, and surface morphology, determines performance, usefulness, and lifetime of the coatings for a given application. A dramatic example of this microstructure–property relationship has been demonstrated by Chun et al. [4–7] who investigated the effect of crystalline texture of TiN layers in microelectronic applications. They found that TiN diffusion barriers with

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (219–227)

111 texture, that is, layers in which the crystalline grains are oriented such that the ⟨111⟩ direction is perpendicular to the substrate surface, exhibit a break down temperature of 450 °C, whereas 001-oriented layers are stable up to 560 °C, a technologically very relevant improvement.

Texture evolution in NaCl structure transition metal nitrides has been studied widely and the results of these investigations provide insight into the atomistic processes that govern layer growth. During the initial stages of growth, primarily 001- and 111-oriented grains nucleate with approximately equal number densities [8, 9]. However, grain growth rates are a function of crystalline orientation. This leads, with increasing film thickness, to the development of strong preferred layer orientation that is either 001 or 111, depending upon film deposition conditions, [4, 8–11]. A model system to study the competition between 111- and 001-oriented grains is the growth of ScN on single crystal MgO(001) substrates. While both layer and substrate crystallize with a cubic structure, there is a 7% lattice mismatch between MgO (a_o = 4.213 Å) and ScN (a_o = 4.501 Å), leading to two possible orientations for ScN nuclei. ScN grains can grow with 001 orientation and a cube-on-cube epitaxial relationship with $(001)_{ScN}||(001)_{MgO}$ and $[100]_{ScN}||[100]_{MgO}$. Alternatively, 111-oriented grains also nucleate with local epitaxy. In that case, strained triangular unit cells of the ScN (111) planes mesh with the squares of the MgO(001) surface, resulting in an epitaxial relationship with $(111)_{ScN}||(001)_{MgO}$ and $[11\bar{2}]_{ScN}||[110]_{MgO}$. Exactly these two grain orientations are observed experimentally by a combination of X-ray diffraction ω-2θ scans and pole figures, as well as electron diffraction patterns from ScN/MgO(001) layers grown in pure N_2 discharges at 750 °C [12]. Cross-sectional transmission electron microscopy shows that these layers exhibit a columnar microstructure. Both grain orientations nucleate with approximately equal volume fractions. However, the width of the 001 columns, which have broad bases and peaked tops, decreases with increasing layer thickness, and they are overtaken by columns with 111 orientation. Thus, the preferred orientation evolves toward a purely 111 texture within ≈40 nm as the 001 grains grow out of existence.

The lowest surface energy face for cubic TM nitrides is the (001), because metal–nitrogen bonds are along perpendicular ⟨100⟩ directions. Hence, 001 is expected to be the preferred orientation under growth conditions where adatom mobilities are sufficiently high to favor crystallites bound by low-energy planes. However, the described result from ScN, together with reports on TiN [8], $Ti_{0.5}Al_{0.5}N$ [13], and TaN [14], all indicate the opposite trend. Thus, the eventual dominance of the 111 versus 001 texture during competitive columnar growth must be due to kinetic limitations during film growth rather than thermodynamic driving forces, as described in the following argument.

An adatom on a (111) surface has three backbonds to N surface atoms and consequently a low potential energy and small surface mobility. In contrast, adatoms on (001) surfaces, with only one N backbond diffuse at much higher rates and may cross a grain boundary to a 111-oriented grain. Thus, adatoms that are stochastically deposited near grain boundaries and, through surface diffusion, sample sites on both sides of the boundary have a higher probability of becoming incorporated at the low-diffusivity (111) surface, which provides the more stable, lower potential energy sites. Conversely, adatoms on high-diffusivity planes have larger mean free paths with correspondingly higher probabilities to move off the plane and become trapped on adjacent grains. The consequence is a net adatom flux from the 001- to the 111-oriented grains, leading to the 111 grains overgrowing the 001 which die out.

This argument is consistent with detailed *ab initio* density functional studies of adsorption and binding energies of adatoms on TiN(001) and TiN(111) [15]. Calculated adsorption energies for Ti adatoms are 3.30 and 10.09 eV, with activation barriers for surface diffusion of 0.35 and 1.74 eV, on TiN(001) and TiN(111), respectively. Thus, Ti adatoms are much more stable and have a much lower surface mobility on 111- versus 001-oriented grains. Atomistic Monte Carlo simulations for face centered cubic metal film growth [16] indicate a trend similar to that reported for the growth of transition metal nitrides. The higher-energy surface emerges as the preferred orientation, primarily due to kinetic limitations.

In summary, the texture of transition metal nitride layers gradually evolves toward 111, due to a combination of a large cation surface diffusion length on the (001) surface combined with a considerably lower adatom potential energy on the (111) versus the (001) surface, leading to a net flux from 001- to 111-oriented grains. However, this effect can be totally reversed by high-flux ion irradiation during growth, as described in the following paragraphs.

Magnetically unbalanced magnetron sputter deposition methods have been developed to investigate plasma–surface interactions during film growth [17]. In such experiments, an external magnetic field is used to shape the plasma near the substrate, resulting in an increase by up to 2 orders of magnitude in the ion flux incident at the growing film with essentially no effect on the sputtered atom flux. The primary advantage of this technique is that it allows independent control over ion energy, determined by the substrate bias, and ion flux. It has been applied to the growth of polycrystalline and single-crystal transition metal nitride layers with dramatic effects on microstructural evolution and texture [8, 9, 13, 14, 18, 19]. For example, TiN layers grown at 350 °C with low ion-to-Ti flux ratios, $J_{N_2^+}/J_{Ti} \leq 1$, and low ion energies $E_i = 20$ eV exhibit 111 texture. However, increasing $J_{N_2^+}/J_{Ti}$ to ≥ 5, with all other deposition parameters remaining constant, results in 001 texture [8, 9], a total reversal. Very similar results are also reported for δ-TaN [14] grown in mixed Ar + 15% N_2 discharges and an ion energy $E_i = 20$ eV. Films grown with high ion-to-Ta flux ratios $J_i/J_{Ta} \geq 7.4$ exhibit a complete 001 texture while those grown with low ion flux ($J_i/J_{Ta} \leq 6.3$) have a strong 111 preferred orientation.

The energy E_i of the ions impinging on the surface during growth of transition metal nitrides, however, strongly determines their effect. Low E_i values ($\leq$15 eV) typically have no noticeable effect, $E_i \approx 20$ eV provides sufficient momentum transfer to alter surface kinetics, whereas larger energies, $E_i \geq 30$ eV, cause bulk defects and compressive stress. In the case of TaN, 8.5 eV is insufficient to observe an effect on texture evolution whereas $E_i = 30$ eV results in nucleation of a secondary hexagonal phase ε-TaN [14]. For ScN,

growth with $E_i = 13$ eV yields layers with mixed grain orientation, texture is controlled using 20 eV ions, and $E_i = 50$ eV causes high compressive stress due to the inclusion of N_2 gas bubbles in the ScN matrix [18]. The onset for ion irradiation-induced bulk defect formation in TiN is reported to occur at 43 eV with $T_s = 650$ °C [20]. Considerable amounts of N_2 bubbles form in TiN at $E_i > 300$ eV [21].

Several models have been proposed to explain the effect of ion irradiation on texture evolution in cubic transition metal nitrides. Pelleg et al. [22] and later Oh and Je [10, 11] proposed that the orientation of polycrystalline TiN films should initially be 001, corresponding to the lowest energy surface [21]. However, ion irradiation causes compressive stress, leading to preferred 111 texture, because the elastic modulus is lower in the [111] direction. The change in orientation in this model is driven by the film/substrate system, minimizing the total free energy.

However, investigations by Greene et al. [8, 9] mentioned above, showed that low-energy (20 eV) N_2^+ ions impinging on the layer surface during reactive TiN deposition have a dramatic effect on texture, without introducing strain. They therefore proposed a model in which texture evolution is determined by competition between thermodynamic and kinetic driving forces. The formation of 111 texture under low-flux conditions was ascribed to limited adatom mobilities on (111) surfaces, leading to larger cation residence times, and consequently, higher cation incorporation probabilities on 111- versus 001-oriented grains. In contrast, high-flux ion irradiation gives rise to increased surface diffusivity due to ion–adatom momentum transfer and therefore has an effect similar to raising the growth temperature. Thus, under high ion flux conditions, texture evolves toward the lowest surface energy, which is the 001 [21].

More recently, a detailed atomistic model has been proposed, based on results from density functional calculations [15]. Texture evolution toward 111 orientation has been ascribed primarily to the lower cation adatom chemical potential on the (111) surface, also related to the low surface mobility on 111 grains, similar to the argument of Greene et al. [8, 9]. However, it was found that the impingement of high-flux low-energy (~20 eV) N_2^+ ions on the layer surface results in the presence of excess atomic N on the (001) surface, due to collisionally induced N_2^+ dissociation. The atomic N reduces the Ti diffusion length, enhances the surface island nucleation rate, and lowers the Ti chemical potential on the (001) surface, leading to preferential growth of 001-oriented grains due to ion irradiation.

The change in texture from 111 to 001, mediated by ion irradiation-induced surface diffusion effects, also has a dramatic effect on layer microstructures. The limited adatom mobility on (111) surfaces of 111-textured layers leads to increased surface roughness and, due to atomic shadowing effects, underdense microstructures. Consequently, 111-oriented layers exhibit underdense columnar microstructures containing open grain boundaries and both inter- and intracolumnar voids. In contrast, the higher cation surface diffusivity on 001-oriented layers, obtained by growth with high ion-to-metal ratios, results in smoother surfaces and fully dense microstructures, which is generally preferred for applications such as hard corrosion-resistant coatings or diffusion barriers.

In summary, during the growth of transition metal nitride layers there are primarily two surfaces of importance, (111) and (001), which exhibit distinctively different growth kinetics. The (111) surfaces are under typical growth conditions fully N terminated [15], leading to low cation adatom potential energies and mobilities. In contrast, cations on (001) surfaces diffuse fast until they reach a step edge, surface island, or grain boundary or form an admolecule with a N adatom. The surface coverage of the latter can be altered by changing the flux of low-energy N_2^+ ion irradiation, which, in turn, allows control over growth kinetics and therefore microstructural evolution. The understanding of this anisotropy between (111) and (001) surfaces, together with kinetic surface roughening and atomic shadowing, is essential to the growth of nanopipes as described in the following section.

3. SURFACE MORPHOLOGICAL EVOLUTION

The development of nonplanar surfaces during transition metal nitride thin film growth results in surface mound structures that cause atomic shadowing which, under limited adatom mobility conditions, leads to the formation of voids, as described above, and more importantly, the formation of nanopipes, as described below. Surface morphological evolution, including surface diffusion, surface island growth and decay kinetics, and surface mound formation have been investigated for TiN(001) single crystal layers using *in-situ* scanning tunneling microscopy (STM) studies [20, 23–29].

Karr et al. [23] were the first to report *in-situ* STM results from TiN(001) surfaces that were grown on MgO(001) by reactive magnetron sputter deposition at 750 °C. Figure 1 shows the surface morphology of layers with thicknesses $t = 25, 50, 100$, and 230 nm. In all images, atomic height steps (0.21 nm) are clearly observable. The surface morphology on a large scale is controlled by a regular array of growth mounds, ~1 nm high, that are separated by a characteristic length d ~100 nm. The sequence of images indicates that both d and the vertical amplitude of the surface roughness A slowly increase with film thickness, indicating growth of the surface mounds in both lateral and vertical directions. The formation of the mounds is due to the asymmetry in the attachment of adatoms at ascending versus descending steps, which destabilizes growth on low miscut surfaces [20], and has been reported to lead to mound formation on both metal and semiconductor surfaces [30–32]. Although this process is well established, the reported roughening rates vary widely, with exponents ranging between 0.16 and ≈1 [23]. The roughening and coarsening rates for TiN(001) layer growth are within the range of rates reported for other materials and follow a power law with the exponent 0.25, that is, $d \propto h^{0.25}$ and $A \propto h^{0.25}$. Increasing the growth temperature does not affect the roughening exponent; however, it slightly increases d and decreases A [20]. Thus, the surface mounds become wider and less high at higher temperature, due to larger adatom diffusion lengths, similar to results from low-temperature growth of Ge(001) [33]. Ion irradiation during growth reduces the effective strength of the growth instability that drives the formation of growth

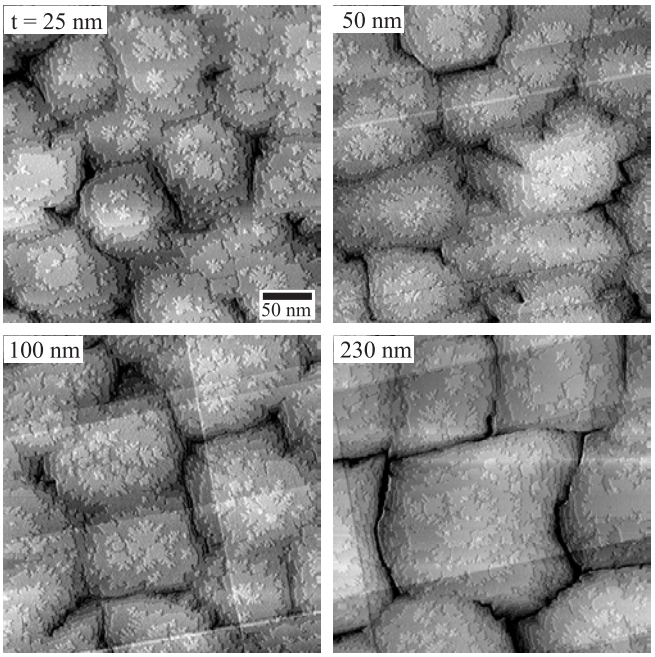


Figure 1. Scanning tunneling microscopy images of TiN(001) films grown on MgO(001) by reactive magnetron sputter deposition at $T =$ 750 °C as a function of film thickness $t = 25$, 50, 100, and 230 nm. Reprinted with permission from [23], B. W. Karr et al., *Appl. Phys. Lett.* 70, 1703 (1997). © 1997, American Institute of Physics.

mounds [20] and also reduces the diffusion length of Ti adatoms and consequently d, due to the presence of atomic N on the surface [15].

A study by Wall et al. [29] focused on surface diffusion and nucleation kinetics to develop a quantitative understanding of the relevant processes determining the surface morphological evolution of TiN(001). In this study, 0.3 monolayers of TiN were deposited on large, atomically smooth TiN(001) terraces. The resulting surface islands, which are one atomic step high, are dendritically shaped and increase in size with increasing temperature. The nucleation length L_n, which is defined as the characteristic separation between these islands, increases from 6 nm at 510 °C to 9 nm at 600 °C to 17 nm at 800 °C. These data, together with a derivation by Venables [34] showing the exponential growth of L_n versus T and the relationship between L_n and the surface diffusion, are used to determine the effective adatom surface diffusion activation energy for TiN(001), found to be $E_s = 1.4 \pm 0.1$ eV [29].

An alternative approach to studying surface diffusion and island kinetics was reported by Kodambaka et al. [25–28], who examined TiN(001) island decay kinetics by STM annealing experiments. The decay of islands in the presence of larger islands is due to coarsening (Ostwald ripening), as described by the Gibbs–Thompson equation showing that the equilibrium adatom concentration around an island increases with decreasing radius r. Small islands have a higher curvature and hence a higher two-dimensional spreading pressure than larger islands, resulting in the decay of the smaller and growth of the larger islands. Kodambaka et al. [26] modeled island decay, adatom diffusion, and adatom density distributions and extracted from the experimental data a value for the activation energy for adatom detachment plus diffusion of 3.4 ± 0.3 eV.

4. ATOMIC SHADOWING

Atomic shadowing, also termed self-shadowing, in the context of thin film growth refers to the part that the surface morphology plays in the capture probability of incident atoms from the gas phase. This is primarily due to surface mounds having a higher probability than surface valleys of capturing incident atoms, leading to an inequality in layer growth rates as a function of location on the surface. The development of surface morphology in the presence of atomic shadowing has been modeled by a variety of theoretical approaches [35–43].

The model by Karunasiri et al. [35] accounts for nonlocal atomic shadowing and surface diffusion and contains a Gaussian white noise term. They found that for small values of the diffusion constant, the growing layer surface develops a self-similar mountain landscape while larger diffusion values result in the growth of compact flat films. Bales and Zangwill [36, 37] showed that for the case of negligible surface diffusion, the maxima (i.e., surface mounds) flatten as growth proceeds while, in contrast, the minima steadily evolve toward grooves. Also, in the absence of diffusion, no characteristic length scale is favored, that is, self-similar starting surfaces result in self-similar final morphology, consistent with "cauliflower"-like morphologies observed experimentally [44]. However, when surface diffusion is present, that is, at growth temperatures that are a significant fraction of the melting point, smoothing of the surface occurs with a characteristic length scale, determined by the diffusion length. This smoothing is in direct competition to the destabilizing shadowing term, which causes groove formation.

The model by Guo and co-workers [38–40] shows that for low values of the angle spread of the incoming deposition flux $\theta_{\max} < 5°$, shadowing effects are minor and the surface simply roughens. However, larger $\theta_{\max}$ values result in columnar structures with deep grooves. Columns also coalesce or compete; for example, small columns are overshadowed and evolve into grooves. This column competition and coarsening follows a power law for which the exponent depends on $\theta_{\max}$, as shown by simulations by Krug and Meakin [41] as well as by Drotar et al. [42]. The combination of atomic shadowing with anisotropic material effects including surface free energies and/or surface diffusion is key to the development of texture evolution during polycrystalline thin film growth, as shown by a model of Karpenko et al. [43].

When the deposition flux occurs from oblique angles, atomic shadowing not only results in surface roughening or the formation of grooves but also can be used in a controlled manner to design well-separated complex three-dimensional nanostructures. This so-called glancing angle deposition (GLAD) technique exploits atomic shadowing to obtain layers with anisotropic microstructure and therefore anisotropic physical properties, including optical activity. This technique has been known for several decades [45, 46] but has gained considerable attention in more recent years due to the wide range of achievable nanostructures [47–70]. Atomic shadowing under limited adatom mobility conditions

results, if the growth occurs at glancing angles ($\alpha \gtrsim 70°$), in a microstructure that is dominated by well-separated columns, which themselves can exhibit complex shapes by virtue of substrate rotation during deposition [49]. The term columnar thin film (CTF) refers to such layers consisting of separated columns whereas sculptured thin films (STFs) are layers for which the columns themselves have an engineered complex structure. The porosity of the CTFs as well as the column tilt angle β increases with increasing deposition angle α. However, the porosity and β can be varied independently using controlled substrate rotation [51]. For example, if substrate rotation is fast relative to the time needed to grow a thickness comparable to the column width, columns grow straight (perpendicular to the substrate surface) although the deposition flux is at glancing angle [52].

A major application for STFs is optically active coatings [53]. Layers with tilted columns exhibit a strong dependence of optical transmission on the angle of incidence and polarization [54]. Biaxial films with large birefringence as well as thin film wave plates have been achieved using the GLAD technique [55, 56]. Layers containing helices act as helicoidal bianisotropic medium [57, 58]. They directly affect circular polarized light. For example, a circular polarization spectral hole filter can be realized by two structurally left-handed chiral STFs with a 90° twist between them [59, 60]. Regular arrays of columns are obtained using initial patterning of the substrate followed by GLAD. Such layers have potential applications as photonic crystals [52, 61, 62]. Columnar thin films also exhibit exceptional magnetic properties with possible applications in magnetic storage devices [63–66].

Shadowing effects during deposition at an oblique angle were also applied to grow Fe nanowire arrays on NaCl(110) templates [68], amorphous Si square spirals on colloid substrates [69], and nanoflowers using a periodic varying substrate rotation speed [70].

Glancing angle deposition was also studied theoretically with both analytical and numerical approaches [71–78]. These studies provided insight into the processes that determine the columnar microstructure of layers grown from oblique angles. There is a general consensus that deposition angle and surface diffusion are key parameters. Also, the column tilt angle β is smaller than the deposition angle α; however, the various studies do not agree on the exact form for the α versus β relationship and also indicate that material parameters like the crystal structure may be important.

5. NANOPIPES

Nanopipes have been first observed in single crystal ScN(001) layers [79], followed by reports on CrN [19], TaN [80], and $Ti_{1-x}W_xN$ [81]. ScN layers grown at $T_s = 800$ °C on TiN(001) buffer layers on MgO(001) exhibit a cube-on-cube epitaxial relationship with the substrate: $(001)_{ScN} \| (001)_{TiN} \| (001)_{MgO}$ and $[100]_{ScN} \| [100]_{TiN} \| [100]_{MgO}$, as determined by a combination of X-ray diffraction and transmission electron microscopy (TEM) analyses [79]. This is in contrast to the results mentioned above, showing that growth of ScN directly on MgO(001) leads to a polycrystalline film exhibiting a mixture of two grain orientations. The epitaxial ScN layers are in a mild state of in-plane compressive strain ($\varepsilon_\| = -0.8\%$), due to differential thermal contraction during cooling from the film growth temperature, because the thermal expansion coefficient for ScN, 4×10^{-6} K^{-1} [82], is much smaller than that for MgO, 1.3×10^{-5} K^{-1} [83].

A typical cross-sectional transmission electron microscopy (XTEM) image with a corresponding selected area electron diffraction pattern, in this case from an epitaxial 82-nm-thick ScN film deposited on a TiN(001) buffer layer, is shown in Figure 2. The 100 (Fig. 2a) zone axis diffraction pattern, obtained with a 0.4 μm aperture centered at the ScN/TiN interface, exhibits two sets of symmetric reflections characteristic of single crystals and separated by a distance consistent with the lattice constant mismatch between ScN and TiN. XTEM tilting experiments reveal an abrupt ScN/TiN interface, with misfit dislocations visible, at the position indicated in Figure 2b. The ScN layer appears defect-free (other than the presence of dislocations) up to a thickness t_c of ≈15 nm. Above the defect-free region, nanopipes with a width of ∼1 nm extend along the growth direction and continue to the free surface where they terminate at cusps. The apparent termination of some nanopipes below the surface in Figure 2b is due to a gradient in the thickness of the XTEM specimen, which is thinnest at the TiN buffer layer and thickest at the ScN free surface. The nanopipes exhibit contrast reversal in under- and overfocus imaging, indicating that they correspond to empty holes in an otherwise fully dense single crystal. The surface is relatively rough, exhibiting a periodic mound structure with an average mound height of ≈7 nm. Between the surface mounds are deep cusps, which connect to the nanopipes.

X-ray photoelectron spectroscopy (XPS) depth profile measurements confirm that the nanopipes are continuously open holes [79]. XPS shows that when the layers are exposed to an air atmosphere, the internal surfaces adsorb oxygen. That is, oxygen enters the layer from the surface through the cusps and adsorbs then at the exposed surfaces of the nanopipes. The nanopipe-free bottom of the layer, however, remains oxygen free.

The experimental results described above show that (a) nanopipes are not due to tensile stress, as could be imagined, because the presented layers exhibit even a slight compressive stress, (b) they do not form at the substrate/layer interface but develop after a certain critical thickness, (c) they

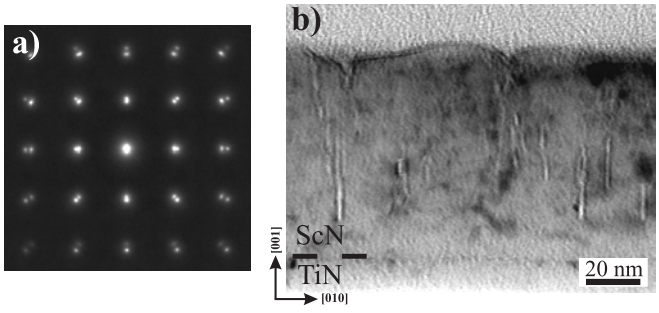


Figure 2. (a) Selected area electron diffraction pattern and (b) corresponding bright-field XTEM micrograph from an epitaxial ScN film grown on TiN/MgO(001). Reprinted with permission from [79], D. Gall et al., *J. Appl. Phys.* 86, 5524 (1999). © 1999, American Institute of Physics.

terminate at surface cusps, and (d) nanopipes provide a continuously open path trough the entire layer with exception of the nanopipe-free bottom. Based on these experimental observations, the following model for the nanopipe formation has been developed.

Initially the surface roughens kinetically to form a periodic mound structure [18], as illustrated in Figure 3a. Such mound structures are typical for films grown at temperatures that are low with respect to the melting point and have been observed in semiconductors [84–88] and metals [89]. The primary origin of kinetic roughening is the presence of Ehrlich–Schwöbel barriers [90] to the migration of adatoms over down-steps and/or deep traps at step edges on growing surfaces. This leads to a divergence in adatom flux and, hence, increased nucleation on terraces. Adatoms have a higher probability to be incorporated at ascending rather than descending step edges, which results in a net adatom flux toward surface mounds, giving rise to surface roughening and faceting during film growth.

When the mounds reach a certain height (Fig. 3b), roughening is exacerbated due to atomic shadowing from surface mounds. Sputter deposition typically results in a cosine distribution of the incoming atom flux with the highest flux at an azimuthal angle of 45° [16]. Atoms reaching the surface at non-normal angles have a higher probability to impinge on mounds than in valleys, leading to a decrease in the local deposition rate in the valleys which results in deep surface cusps (Fig. 3c). The surface around a cusp contains a high density of step edges or may even form (111) facets. Consequently, the adatom mobility around surface cusps is relatively low, as discussed above when diffusion on (001) and (111) surfaces was compared. Thus, surface diffusion is insufficient for filling of the trenches, leading to the formation of nanopipes as shown in Figure 3d.

The observation that nanopipes form only above a critical film thickness ($t > t_c \simeq 15$ nm) indicates that surface roughening occurs gradually during growth and a critical mound height, which depends on both the angular incident flux distribution and the adatom mobility, is required for maintaining continuously open nanopipe structures. Nanopipes elongate along the growth direction because the average direction of the incoming atom flux is normal to the layer surface. This directionality is drastically enhanced by the high adatom mobility on {100} surfaces, which causes nanopipes to be preferentially bound by {100} planes, as observed using plan-view transmission electron micrographs. An example is shown in Figure 4, taken from a

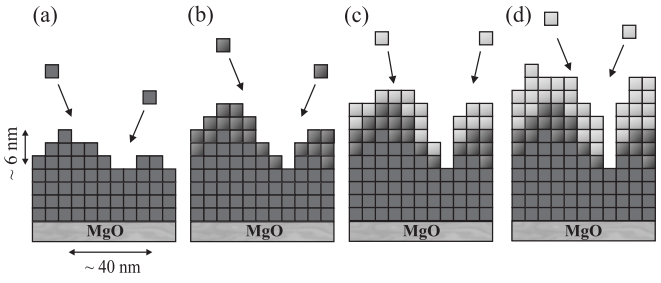


Figure 3. Schematic of nanopipe formation in single crystal transition metal nitride layers.

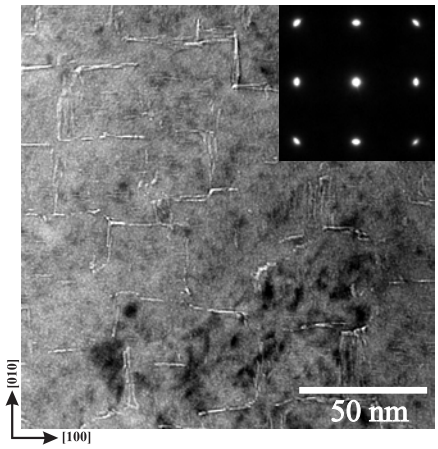


Figure 4. Plan-view TEM micrograph and corresponding diffraction pattern from a ScN(001)/MgO(001) layer showing the arrays of rectangular shaped nanopipes aligned along orthogonal ⟨100⟩ directions. Reprinted with permission from [18], D. Gall et al., *J. Appl. Phys.* 84, 6034 (1998). © 1998, American Institute of Physics.

ScN(001)/MgO(001) layer grown at 750 °C with $J_i/J_{Sc} = 14$ and $E_i = 20$ eV [18]. This plan-view sample was prepared by step-wise ion milling through the film from both sides. Thus, the micrograph is a plan-view from the center of the 345-nm-thick layer, showing a cross-section through the nanopipes, which are elongated along the growth direction perpendicular to the image plane. The nanopipes are rectangular in cross-section with a width of ~1 nm by 1–15 nm along orthogonal ⟨100⟩ directions. They are bound by high adatom mobility {100} planes. Their rectangular elongated shape is explained by their formation; approximately square-shaped surface mounds, with edges along low-energy ⟨100⟩ directions (i.e., along Sc-N bonding directions), adjoin along ⟨100⟩ with the development of valleys, which are consequently also elongated along ⟨100⟩. These transform, with increasing layer thickness, to deep cusps and ultimately nanopipes as shown in Figure 3. The square-shaped surface mounds, which have an average mound-to-mound spacing of 30 ± 5 nm, are centered between the nanopipes and their locations and shapes can easily be imagined when looking at Figure 4. Because the nanopipes are directly linked to the periodic surface roughness, they form a self-organized periodic array with, however, only very limited long-range order. Long-range order and complex nanostructures may nevertheless be obtained by initial patterning and controlled surface morphological evolution, as indicated in the conclusion/outlook section.

The nanopipe density can be controlled by ion irradiation during growth. This has been reported in a study on epitaxial CrN layers [19], which combined atomic force microscopy (AFM) and TEM to investigate the correlation of surface morphological evolution with nanopipe formation. Epitaxial CrN(001) layers of 0.5 μm thickness were grown at $T_s = 600$ °C by ultrahigh vacuum magnetron sputter deposition on MgO(001) substrates. Growth under typical conditions for magnetron sputtering, with an ion energy E_i of 12 eV and an ion-to-metal ratio $J_{N_2^+}/J_{Cr} = 1.7$, leads to a surface morphology that exhibits square-shaped growth mounds with edges predominantly aligned along low-energy

⟨100⟩ directions. Lateral mound sizes range from 130 to 300 nm with a surface width $\langle w \rangle$ of 1.80 ± 0.04 nm, an average peak-to-valley mound height $\langle h \rangle = 5.09 \pm 0.11$ nm, and an in-plane correlation length $\langle d \rangle$ of 208 ± 20 nm. However, growth at a high ion flux, $J_{N_2^+}/J_{Cr} = 14$, with all the other deposition parameters remaining, results in a strikingly different surface morphology. The growth mounds are much larger and elongated along ⟨110⟩, while they are dendritic in the orthogonal direction. Both the dendritic shape and the elongation along ⟨110⟩ indicate large surface diffusion lengths and a growth instability driven by the preferred adatom attachment at ascending versus descending step edges. The average mound height $\langle h \rangle = 2.52 \pm 0.03$ nm of samples grown with high ion irradiation flux is ∼2 times smaller than that for the $J_{N_2^+}/J_{Cr} = 1.7$ layer whereas the in-plane correlation length $\langle d \rangle = 520 \pm 100$ nm is 2.5 times larger. This results in a decrease of the average aspect ratio $\langle h \rangle / \langle d \rangle$ by a factor of 5 when $J_{N_2^+}/J_{Cr}$ is increased from 1.7 to 14.

These results from epitaxial CrN(001) layers show that increasing the incident ion-to-metal flux ratio during growth, using low N_2^+ ion energies, results in smoother surfaces with both smaller mound heights and decreased mound aspect ratios [19]. High-flux, low-energy, ion irradiation leads to higher adatom mean surface diffusion lengths and, consequently, larger surface features. In addition to larger adatom diffusion lengths on terraces, momentum transfer from low-energy ion irradiation increases the probability of adatoms crossing descending step edges. This contributes to the filling of trenches between growth mounds, thereby reducing atomic shadowing. The surface smoothing obtained at higher ion-to-metal flux ratios has important consequences on the formation of nanopipes, as described in the following.

XTEM investigations [19] show that CrN layers grown with $J_{N_2^+}/J_{Cr} = 1.7$ exhibit 1–5-nm-wide nanopipes along the [001] growth direction. The average lateral separation between nanopipes is 180 ± 40 nm, approximately equal to the growth mound in-plane correlation length $\langle d \rangle = 208 \pm 20$ nm determined by AFM. The nanopipes extend along the growth direction to the free surface where they terminate at cusps between growth mounds, suggesting, as discussed above, that the formation of nanopipes is directly related to the presence of the surface cusps between periodic growth mounds that cause strong atomic shadowing.

CrN(001) layers grown with $J_{N_2^+}/J_{Cr} = 6$ show a considerably different nanopipe distribution: The layers are defect-free (no nanopipes) over a thickness $t^* \simeq 250$ nm and exhibit a low nanopipe density at $t > t^*$. The nanopipes, as is the case for the layer grown at the low ion-to-metal ratio, have widths ranging from 1 to 5 nm. However, they are not continuous and appear and disappear along the growth direction. At even higher ion-to-metal ratios, $J_{N_2^+}/J_{Cr} = 10$ and 14, the nanopipe formation is nearly completely suppressed.

Therefore, the strong effects of low-energy N_2^+ ion irradiation during deposition on the surface morphology of CrN(001) layers also control the formation rate of nanopipes. Layers grown with low ion-to-metal ratios, $J_{N_2^+}/J_{Cr} = 1.7$, exhibit rough surfaces composed of growth mounds with relatively high aspect ratios and short in-plane coherence lengths which, within the experimental uncertainty, are equal to the measured average separations between nanopipes. For CrN(001) layers grown with medium ion-to-metal ratios, $J_{N_2^+}/J_{Cr} = 6$, the in-plane correlation length is larger, corresponding to a lower mound number density. This, in turn, results in a reduction in the nanopipe number density by more than a factor of 3 compared with CrN(001) layers grown with $J_{N_2^+}/J_{Cr} = 1.7$. In addition, nanopipes form only above a critical film thickness ($t^* \geq 250$ nm) and are found to be discontinuous along the growth direction. The rate of surface roughening occurs more gradually during growth with $J_{N_2^+}/J_{Cr} = 6$. Thus, at $t < t^*$, the mound aspect ratio is too small to provide sufficient atomic shadowing to form cusps and, hence, nanopipes. Even at $t > t^*$, the surface roughness is not adequate to sustain continuous nanopipe formation, leading to discontinuous nanopipes. At high ion-to-metal flux ratios, $J_{N_2^+}/J_{Cr} \geq 10$, layer surfaces are relatively smooth such that shadowing is negligible and essentially no nanopipes are formed.

6. CONCLUSIONS

Nanopipes in transition metal nitrides form due to atomic shadowing from a surface mound structure under limited adatom mobility conditions, as described in detail above. Nanopipe density, average spacing, and length can be controlled by thin film deposition parameters, including growth temperature, N_2 partial pressure, and ion irradiation energy and flux. This is due to changes in the surface morphology, which is primarily determined by surface diffusion and reaction processes on both (001) and (111). The surface mounds are roughly square shaped with edges along perpendicular ⟨100⟩ directions, leading to cusps, elongated along ⟨100⟩ directions, and the formation of nanopipes with rectangular cross-sections that are aligned along the [001] growth direction. The nanopipe width, 1 ± 0.5 nm, is determined by the angular distribution of the deposition flux, which controls the degree of shadowing, whereas the orthogonal elongation is defined by the width of the surface mounds. The average inter-nanopipe separation is equal to the in-plane correlation length and is therefore fully determined by the surface morphology.

While deposition parameters control the average distance between nanopipes by virtue of affecting surface diffusion processes and thus surface morphology, no long-range order of nanopipes has been achieved yet. However, various schemes could be envisioned that would provide the means for controlled nanopipe positioning. One approach would be to pattern the substrate before growth. The epitaxial transition metal nitride layer would then yield nanopipes that are positioned at the grooves of the initial substrate patterning. Although the patterning could be done with resolutions of ∼50 nm, the self-limiting shadowing processes would result in 1-nm-wide nanopipes. This corresponds to a self-organized feature size reduction by a factor of 50.

Further development of this technique could be envisioned. In particular, it should be possible to control the nanopipe width by glancing angle deposition techniques that directly affect the degree of atomic shadowing. In that way, arrays of nanopipes that exhibit a desired width can be synthesized. The width could even be varied as a function of layer thickness by simply changing the incidence angle of

the deposition flux. For example, if the deposition angle is increased for a short time during growth, the nanopipe width will be larger at this given layer height, leading to controlled in-plane nanochannel formation.

These advanced nanopipe-based structures have potential applications as templates for tunneling devices, sensors, and two-dimensional conductors. The nanochannels may be used for molecular transport and single molecule chemistry. Periodic arrays will be useful in optoelectronics and photonic materials. Other potential applications include magnetic storage, large surface area devices, thermal barriers, and optical coatings.

GLOSSARY

Ab initio [Calculations] performed using only fundamental physical constants, that is, no experimental data or empirical numbers.

Atomic shadowing The effect that the surface morphology plays in the capture probability of incident atoms.

Epitaxy/epitaxial growth Thin film growth where the crystallographic orientation of the grown layer aligns with that of the substrate.

Glancing angle deposition (GLAD) Exploits atomic shadowing effects to create layers with columnar structures.

REFERENCES

1. L. E. Toth, "Transition Metal Carbides and Nitrides." Academic, New York, 1971.
2. V. A. Gubanov, A. L. Ivanovsky, and V. P. Zhukov, "Electronic Structure of Refractory Carbides and Nitrides." Cambridge Univ. Press, Cambridge, UK, 1994.
3. K. Schwarz, *Crit. Rev. Solid State Mater. Sci.* 13, 211 (1987).
4. J.-S. Chun, J. R. A. Carlsson, P. Desjardins, D. B. Bergstrom, I. Petrov, J. E. Greene, C. Lavoie, C. Cabral, Jr., and L. Hultman, *J. Vac. Sci. Technol., A* 19, 182 (2001).
5. J.-S. Chun, I. Petrov, and J. E. Greene, *J. Appl. Phys.* 86, 3633 (1999).
6. J.-S. Chun, P. Desjardins, C. Lavoie, I. Petrov, C. Cabral, Jr., and J. E. Greene, *J. Vac. Sci. Technol., A* 19, 2207 (2001).
7. J.-S. Chun, P. Desjardins, C. Lavoie, C.-S. Chin, C. Cabral, Jr., I. Petrov, and J. E. Greene, *J. Vac. Sci. Technol., A* 19, 2207 (2001).
8. J. E. Greene, J.-E. Sundgren, L. Hultman, I. Petrov, and D. B. Bergstrom, *Appl. Phys. Lett.* 67, 2928 (1995).
9. L. Hultman, J.-E. Sundgren, J. E. Greene, D. B. Bergstrom, and I. Petrov, *J. Appl. Phys.* 78, 5395 (1995).
10. U. C. Oh and J. H. Je, *J. Appl. Phys.* 74, 1692 (1993).
11. J. H. Je, D. Y. Noh, H. K. Kim, and K. S. Liang, *J. Appl. Phys.* 81, 6126 (1997).
12. D. Gall, I. Petrov, L. D. Madsen, J.-E. Sundgren, and J. E. Greene, *J. Vac. Sci. Technol., A* 16, 2411 (1998).
13. F. Adibi, I. Petrov, J. E. Greene, L. Hultman, and J.-E. Sundgren, *J. Appl. Phys.* 73, 8580 (1993).
14. C.-S. Shin, D. Gall, Y.-W. Kim, N. Hellgren, I. Petrov, and J. E. Greene, *J. Appl. Phys.* 92, 5084 (2002).
15. D. Gall, S. Kodambaka, M. A. Wall, I. Petrov, and J. E. Greene, *J. Appl. Phys.* 93, 9086 (2003).
16. H. Huang, G. H. Gilmer, and T. D. de la Rubia, *J. Appl. Phys.* 84, 3636 (1998).
17. I. Petrov, F. Adibi, J. E. Greene, W. D. Sproul, and W.-D. Münz, *J. Vac. Sci. Technol., A* 10, 3283 (1992).
18. D. Gall, I. Petrov, N. Hellgren, L. Hultman, J.-E. Sundgren, and J. E. Greene, *J. Appl. Phys.* 84, 6034 (1998).
19. D. Gall, C.-S. Shin, T. Spila, M. Odén, M. J. H. Senna, J. E. Greene, and I. Petrov, *J. Appl. Phys.* 91, 3589 (2002).
20. B. W. Karr, I. Petrov, D. G. Cahill, and J. E. Greene, *Phys. Rev. B* 61, 16137 (2000).
21. L. Hultman, J.-E. Sundgren, and J. E. Greene, *J. Appl. Phys.* 66, 536 (1989).
22. J. Pelleg, L. Z. Zevin, and S. Lungo, *Thin Solid Films* 197, 117 (1991).
23. B. W. Karr, I. Petrov, D. G. Cahill, and J. E. Greene, *Appl. Phys. Lett.* 70, 1703 (1997).
24. B. W. Karr, I. Petrov, P. Desjardins, D. G. Cahill, and J. E. Greene, *Surf. Coat. Technol.* 94–95, 403 (1997).
25. S. Kodambaka, V. Petrova, A. Vailionis, P. Desjardins, I. Petrov, D. G. Cahill, and J. E. Greene, *Surf. Rev. Lett.* 7, 589 (2000).
26. S. Kodambaka, V. Petrova, S. V. Khare, D. D. Johnson, I. Petrov, and J. E. Greene, *Phys. Rev. Lett.* 88, 146101 (2002).
27. S. Kodambaka, V. Petrova, S. V. Khare, D. Gall, A. Rockett, I. Petrov, and J. E. Greene, *Phys. Rev. Lett.* 89, 176102 (2002).
28. S. Kodambaka, V. Petrova, P. Desjardins, I. Petrov, D. G. Cahill, and J. E. Greene, *Thin Solid Films* 392, 164 (2001).
29. M. A. Wall, D. G. Cahill, I. Petrov, D. Gall, and J. E. Greene, unpublished.
30. J. E. Nostrand, S. Jay Chey, M.-A. Hasan, D. G. Cahill, and J. E. Greene, *Phys. Rev. Lett.* 74, 1127 (1995).
31. H.-J. Ernst, F. Fabre, R. Folkerts, and J. Lapujoulade, *Phys. Rev. Lett.* 72, 112 (1994).
32. D. Johnson, C. Orme, A. W. Hunt, D. Graff, J. Sudijono, L. M. Sander, and B. G. Orr, *Phys. Rev. Lett.* 75, 116 (1994).
33. J. E. Nostrand, S. J. Chey, and D. G. Cahill, *Phys. Rev. B* 57, 12536 (1999).
34. J. A. Venables, *Philos. Mag.* 27, 693–738 (1973).
35. R. P. U. Karunasiri, R. Bruinsma, and J. Rudnick, *Phys. Rev. Lett.* 62, 788 (1989).
36. G. S. Bales and A. Zangwill, *Phys. Rev. Lett.* 63, 692 (1989).
37. G. S. Bales and A. Zangwill, *J. Vac. Sci. Technol., A* 9, 145 (1991).
38. C. Roland and H. Guo, *Phys. Rev. Lett.* 16, 2104 (1991).
39. J. H. Yao, C. Roland, and H. Guo, *Phys. Rev. A* 45, 3903 (1992).
40. J. H. Yao and H. Guo, *Phys. Rev. E* 47 1007 (1993).
41. J. Krug and P. Meakin, *Phys. Rev. E* 47, 17 (1993).
42. J. T. Drotar, Y.-P. Zhao, T. M. Lu, and G.-C. Wang, *Phys. Rev. B* 62, 2118 (2000).
43. O. P. Karpenko, J. C. Bilello, and S. M. Yalisove, *J. Appl. Phys.* 82, 1397 (1997).
44. R. Messier and J. E. Yehoda, *J. Appl. Phys.* 58, 3739 (1985).
45. N. O. Young and J. Kowal, *Nature* (*London*) 183, 104 (1959).
46. D. O. Smith, M. S. Cohen, and G. P. Weiss, *J. Appl. Phys.* 31, 1755 (1960).
47. K. Robbie, M. J. Brett, and A. Lakhtakia, *Nature* (*London*) 384, 616 (1996).
48. K. Robbie, D. J. Broer, and M. J. Brett, *Nature* (*London*) 399, 764 (1999).
49. R. Messier, V. C. Venugopal, and P. D. Sunal, *J. Vac. Sci. Technol., A* 18, 1528 (2000).
50. K. Robbie and M. J. Brett, *J. Vac. Sci. Technol., A* 15, 1460 (1997).
51. K. Robbie, J. C. Sit, and M. J. Brett, *J. Vac. Sci. Technol., B* 16, 1115 (1998).
52. M. Malac and R. F. Egerton, *J. Vac. Sci. Technol., A* 19, 158 (2001).
53. A. Lakhtakia, *Mater. Sci. Eng., C* 19, 427 (2002).
54. M. Suzuki and Y. Taga, *J. Appl. Phys.* 71, 2848 (1992).
55. I. J. Hodgkinson and Q. H. Wu, *Appl. Opt.* 38, 3621 (1999).
56. I. J. Hodgkinson and Q. H. Wu, *Opt. Eng.* 37, 2630 (1998).
57. K. Robbie, M. J. Brett, and A. Lakhtakia, *J. Vac. Sci. Technol., A* 13, 2991 (1995).
58. R. M. A. Azzam, *Appl. Phys. Lett.* 61, 3118 (1992).
59. A. Lakhtakia and R. Messier, *Opt. Photon. News* Sept. 1, 27 (2001).

60. I. J. Hodgkinson, Q. H. Wu, K. E. Thorn, A. Lakhtakia, and M. W. McCall, *Opt. Commun.* 184, 57 (2000).
61. M. Malac, R. F. Egerton, M. J. Brett, and B. Dick, *J. Vac. Sci. Technol., B* 17, 2671 (1999).
62. M. Malac and R. F. Egerton, *Nanotechnology* 12, 11 (2001).
63. A. Lisfi and J. C. Lodder, *Phys. Rev. B* 63, 174441 (2001).
64. R. D. McMichael, C. G. Lee, J. E. Bonevich, P. J. Chen, W. Miller, and W. F. Egelhoff, Jr., *J. Appl. Phys.* 88, 3561 (2000).
65. B. Dick, M. J. Brett, T. J. Smy, M. R. Freeman, M. Malac, and R. F. Egerton, *J. Vac. Sci. Technol., A* 18, 1838 (2000).
66. A. Hagemeyer, H. J. Richter, H. Hibst, V. Maier, and L. Marosi, *Thin Solid Films* 230, 199 (1993).
67. K. Starbova, J. Dikova, and N. Starbov, *J. Non-Cryst. Solids* 210, 261 (1997).
68. A. Sugawara, T. Coyle, G. G. Hembree, and M. R. Scheinfein, *Appl. Phys. Lett.* 70, 1043 (1997).
69. Y.-P. Zhao, D.-X. Ye, P.-I. Wang, G.-C. Wang, and T.-M. Lu, *Int. J. Nanosci.* 1, 87 (2002).
70. Y.-P. Zhao, D.-X. Ye, G.-C. Wang, and T.-M. Lu, *Nano Lett.* 2, 351 (2002).
71. A. G. Dirks and H. J. Leamy, *Thin Solid Films* 47, 219 (1977).
72. S. Lichter and J. Chen, *Phys. Rev. Lett.* 56, 1396 (1986).
73. P. Meakin and J. Krug, *Phys. Rev. A* 46, 3390 (1992).
74. J. Krug and P. Meakin, *Phys. Rev. A* 43, 900 (1991).
75. R. N. Tait, T. Smy, and M. J. Brett, *Thin Solid Films* 226, 196 (1993).
76. Paritosh and D. J. Srolovitz, *J. Appl. Phys.* 91, 1963 (2002).
77. I. Hodgkinson, Q. H. Wu, and A. McPhun, *J. Vac. Sci. Technol., B* 16, 2811 (1998).
78. T. Smy, D. Vick, M. J. Brett, S. K. Dew, A. T. Wu, J. C. Sit, and K. D. Harris, *J. Vac. Sci. Technol., A* 18, 2507 (2000).
79. D. Gall, I. Petrov, P. Desjardins, and J. E. Greene, *J. Appl. Phys.* 86, 5524 (1999).
80. C.-S. Shin, Y.-W. Kim, N. Hellgren, D. Gall, I. Petrov, and J. E. Greene, *J. Vac. Sci. Technol., A* 20, 2007 (2002).
81. F. Tian, J. D'Arcy-Gall, T.-Y. Lee, M. Sardela, D. Gall, I. Petrov, and J. E. Greene, *J. Vac. Sci. Technol.* 21, 140 (2003).
82. B. Hajek, V. Brozek, and H. Duvigneaud, *J. Less-Common Met.* 33, 385 (1973).
83. H. Landolt and R. Börnstein, "Numerical Data and Functional Relationships in Science and Technololgy," *Group III*, Vol. 7, Pt. b1, p. 27. Springer-Verlag, Berlin, 1975.
84. N.-E. Lee, D. G. Cahill, and J. E. Greene, *J. Appl. Phys.* 80, 2199 (1996).
85. G. Xue, H. Z. Xiao, M.-A. Hasan, J. E. Greene, and H. K. Birnbaum, *J. Appl. Phys.* 74, 2512 (1993).
86. N.-E. Lee, G. A. Tomasch, and J. E. Greene, *Appl. Phys. Lett.* 65, 3236 (1994).
87. N.-E. Lee, G. Xue, and J. E. Greene, *J. Appl. Phys.* 80, 2199 (1996).
88. D. J. Eaglesham, H.-J. Gossmann, and M. Cerullo, *Phys. Rev. Lett.* 65, 1227 (1990).
89. J. A. Stroscio, D. T. Piercee, M. D. Stiles, A. Zangwill, and L. M. Sander, *Phys. Rev. Lett.* 75, 4246 (1995).
90. S. C. Wang and G. Ehrlich, *Phys. Rev. Lett.* 70, 41 (1993), 71, 4177 (1993); G. Ehrlich, *Surf. Sci.* 331/333, 865 (1995); A. Gölzhäuser and G. Ehrlich, *Phys. Rev. Lett.* 77, 1334 (1996).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanopore Analysis of DNA

David W. Deamer, Stephen Winters-Hilt

University of California, Santa Cruz, California, USA

CONTENTS

1. INTRODUCTION

A nanopore is defined as a pore having dimensions in the range of 1–100 nm. A nanopore therefore admits particles of molecular size but excludes micrometer-sized particles such as bacterial cells. If the pore is produced in a very thin membrane, such as a lipid bilayer 5 nm thick, it becomes possible to detect polymeric solutes as individual molecules. A voltage is first established across the membrane, so that an ionic current (typically potassium and chloride in 1.0 M salt) is driven through the pore. Larger ionic solutes such as nucleic acid molecules can also be transported through the pore by electrophoresis. When a sufficiently large solute molecule enters the pore, a measurable decrease in ionic current (referred to here as a blockade) is produced that provides information about the molecule. The signal is captured and digitized, after which it can be computationally processed to characterize the particle. Important applications for nanopore analysis of polymers in solution include determinations of concentration, identity, and composition of the solute molecules.

The concept of using small pores to analyze particles of biological relevance was pioneered by Coulter [1] who showed that blood cells in saline suspensions could be detected as they were moved through a pore 100 μm in diameter by a pressure gradient. As each cell enters the pore it changes the resistance to the ionic current that passes through the pore. The change is related to the volume of the cell by the equation $\Delta g = 2\pi r^3 \sigma / L^2$, where r is the radius of the particle (with the volume of a spherical particle being a function of r^3), L is the length of the pore, and σ is the conductivity of the solution [2]. The resulting signals provide information about the cell number and cell volume. In a blood sample, cells such as red cells and leukocytes have different volumes, so that differential counts can be rapidly carried out. This information is useful in diagnosis of blood disorders, and commercial versions of the Coulter counter remain an important instrument in clinical laboratories.

Following the success of pore-based cell counters as analytical tools, interest turned to smaller pores that could be used to detect particles such as biological polymers. However, it is much more difficult to produce a pore in the nanoscopic size range, 2–3 orders of magnitude smaller than used by Coulter. Bean and co-workers [2–4] first used etched nuclear tracks for this purpose. When a high energy nucleus passes through a solid it typically leaves a "track" in which the structure has been chemically altered at the atomic level of structure. This track can later be etched by applying solvents that preferentially dissolve the damaged structures. Tracks are produced by natural radioactive decay in a substance, such as fission of a uranium nucleus present in silicate minerals, or by exposure to a radioactive source such as californium (Cf_{252}) that undergoes spontaneous fission. In the latter case the fragments can be collimated so that the resulting tracks are parallel. The tracks are then etched by exposure to solvents such as hydrofluoric acid or concentrated sodium hydroxide. (Nucleopore filters with pore sizes ranging from 50 nm to several micrometers are produced by

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (229–235)

this process and are commercially available.) If mica etched with 49% HF is used, the pore size at initial breakthrough is approximately 6 nm, which then increases to micrometer size ranges after 30 minutes of etching. DeBlois et al. [4] showed that mica nanopores were sufficiently sensitive so that particles as small as viruses could be detected by the change in resistance as the particle was driven through the pore.

Although mica and polycarbonate films produce pores of controlled size, the pores are in fact tunnels through the relatively thick materials. This limits their resolution, because a globular particle spends time negotiating a tunnel of varying dimensions. Ideally a nanopore will be present in a film that is no thicker than the analyte to be measured. This is particularly true of linear polymers such as nucleic acids, in which the desired resolution is at the level of a single nucleotide.

2. BIOLOGICAL NANOPORES

In the 1970s, it became apparent that biological membranes of cells incorporate nanoscopic channels composed of proteins that are embedded in lipid bilayers. The function of the channels is to provide a gated pore that allows solutes such as sodium, potassium, and calcium ions to cross the otherwise impermeable lipid bilayer barrier that surrounds all cells. Artificial bilayers are readily produced by self-assembly of amphiphilic compounds such as phospholipids. For instance, bilayers of phospholipid 5 nm thick can be produced as membranes covering a 100 micrometer diameter hole in a thin plastic sheet. Such membranes have dimensions appropriate for nanopore analysis, provided that a pore can be produced in the bilayer.

Hladky and Haydon [5] showed that the bacterial antibiotic gramicidin spontaneously inserted into lipid bilayers and could form continuous channels that conducted ionic current. Gramicidin, with a pore diameter of ~0.4 nm, is the smallest useful nanopore. The gramicin pore contains a chain of approximately 10 water molecules that form a single hydrogen-bonded strand across the bilayer. In the presence of a voltage, the pore conducts an ionic current of a few picoamperes, the amplitude depending on the voltage and ion concentration. The gramicidin pore is too small to be used directly to detect solutes other than ions, but Cornell and co-workers [6] have shown that it can be used as a biosensor if it is coupled to a specific antibody. When a target antigen binds to the antibody, the ability of the gramicidin pore to conduct ionic current is inhibited. The reduction in current can then be used to estimate the presence and concentration of the antigen.

The fact that such pores could conduct polar and ionic solutes across a membrane suggested that a larger pore might be used to detect macromolecules in solution. This concept was first tested by Bezrukov et al. [7] who showed that ionic current through the alpha hemolysin pore was affected by the presence of polyethylene glycol (PEG). The current fluctuations appeared only at certain molecular size ranges: PEG molecules that were smaller or larger than the pore had little effect, while PEG in the size range of the pore (0.5–1.5 nm) produced large current fluctuations. This allowed the authors to conclude that PEG could enter the pore and affect ionic current. Because PEG has no net electrical charge, the molecules enter the pore by diffusion, rather than by an energy-dependent process. (See Bezrukov [8] for review.)

At about the same time, Walker and co-workers [9] demonstrated that the hemolysin pore could be modified to serve as a metal biosensor. In this work, the hemolysin was modified by introducing histidine for amino acids 130–134 in the central glycine-rich loop of the protein. The modified pore had a conductance that could be modulated by micromolar concentrations of ions such as zinc, suggesting that modified protein pores such as hemolysin could serve as sensitive metal ion sensors.

3. NANOPORES AS LINEAR POLYMER SENSORS

The next advance in nanopore analysis was to take advantage of the fact that larger ionized solutes could potentially be driven through the pore by an electrical field. This would be particularly important for analyzing linear polymeric solutes such as nucleic acids. The pore diameter permits nucleic acids to be translocated as single strands, rather than duplexes, so that a given strand passes through the pore in strict linear sequence. This fact led to the concept of nanopore sequencing, which depends on the possibility that each base in a nucleic acid strand will modulate the signal in a specific and measurable way. If so, the base sequence in a nucleic acid molecule could be determined at rates far exceeding those now possible.

The hemolysin toxin produced by *S. aureus* is a suitable nanopore because it self-assembles in lipid bilayers to form aqueous pores with diameters just large enough to translocate the nucleotides in a single strand of DNA or RNA (Fig. 1).The entrance to the head of the channel has a diameter of approximately 2.6 nm that contains a ring of lysine residues. The entry opens into a vestibule with an interior diameter of 3.6 nm that in turn leads to the pore's stem which penetrates the lipid bilayer. The average inside diameter of the stem is about 2.0 nm with a 1.5 nm constriction between the vestibule and the stem. This limiting constriction is composed of alternating lysine and glutamate residues, and the remainder of the stem is lined primarily with neutral residues except for a single hydrophobic ring of exposed leucine residues. The opening at the other end of the pore has a 2.2 nm ring of alternating lysine and aspartate residues.

To produce a nanopore, α-hemolysin subunits are introduced into a buffered solution that is in contact with a lipid bilayer which separates the solution into two compartments, designated *cis* and *trans*. This apparatus, in conjunction with a standard amplifier and software, provides rapid response times on the microsecond time scale and substantially lower noise than the larger planar lipid membrane commonly used in single channel research. The open channel of a hemolysin pore carries an ionic current of approximately 120 picoamps (pA) with an applied voltage of 120 millivolts (mV).

Kasianowicz et al. [10] first reported that addition of single-stranded nucleic acid homopolymers produced large numbers of transient ionic current blockades, suggesting that individual molecules were being translocated through the pore and blocking ionic current. Formal proof for this

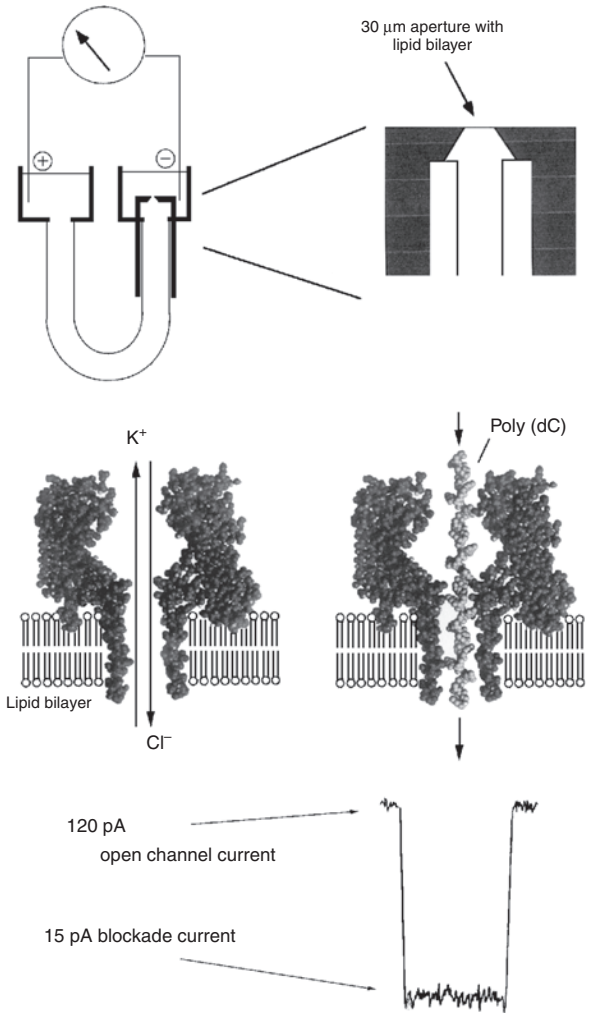


Figure 1. α-Hemolysin pore inserted into a lipid bilayer. The heptameric pore is illustrated in cross section, and a single strand of poly(dC) synthetic DNA is shown translocating from the top through the nanoscopic pore in the stem that penetrates the bilayer. The resulting ionic current blockade is shown at the bottom, with a duration of approximately 200 microseconds for a 100mer.

conjecture followed when it was demonstrated that the number of blockades was correlated with the actual number of single-stranded DNA molecules translocated through the pore [10]. The experimental approach involved a comparison of the ability of DNA as single-stranded (ssDNA) and double-stranded (dsDNA) molecules to pass through the pore. Quantitative PCR was then used to make a direct measurement of nucleic acids appearing on the *trans* side of the membrane, and the numbers were compared with the number of blockade events that occurred during the same time interval. Only ssDNA appeared on the *trans* side, and the ratio of single-stranded DNA molecules appearing on the *trans* side to the number of blockade events was close to 1. This observation permitted the conclusion that single-stranded nucleic acids were in fact translocated through the hemolysin pore, and that the resulting ionic current blockades could be used to gather information about single molecules.

4. BLOCKADE MECHANISM

The process by which an ionic current blockade is produced must entail three phases: capture, entry, and translocation. When a voltage of 120 mV is imposed on the pore, the resulting ionic current consists of potassium and chloride ions. If single stranded nucleic acids are present, a given molecule will occasionally diffuse into the small volume near the mouth of the pore. When the polymer enters the biased region, three outcomes are possible. The first is that the diffusing nucleic acid molecule simply collides with the pore mouth. During this collision, the ionic current may be interrupted for a few microseconds, but the duration and extent of this transient blockade is variable and largely independent of chain length.

The second possibility is that one end of the molecule partially enters the vestibule and remains for periods ranging from tens of microseconds to several milliseconds, after which it either diffuses out of the pore or is drawn completely into the pore stem. While the molecule occupies the vestibule it causes a partial blockade having a characteristic amplitude about half that of the full blockade.

The third possibility is that one end of the nucleic acid is drawn completely into the pore stem, where it produces a full blockade with a duration that is a function of chain length. Translocation occurs as a result of electrophoretic force acting on the anionic phosphate groups of the chain. The pore stem is only 5 nm long, but the nucleic acid strand can be up to thousands of nucleotides in length. Experiments with very short polynucleotides show that the electric field acts primarily on the 10–14 phosphate groups within the stem [11]. The resultant force of about 20–40 piconewtons is sufficient to drive the molecule through the pore with velocities that depend on the composition and secondary structure of the strand. Given the small aperture of the pore (1.5 nm), the polynucleotides must move as an extended linear polyanion. Meller et al. [11, 12] have undertaken detailed studies in which temperature effects on translocation are considered, as well as direct estimates of force acting on a DNA strand during translocation. (See Muthukumar [13] and Lubensky and Nelson [14] for discussions of the physics of polymer translocation through nanopores.)

5. PARAMETERS OF IONIC CURRENT BLOCKADES

The structural properties of the hemolysin nanopore shown in Figure 1 are relevant to understanding the mechanism of ionic current blockades. Here we will assume that the major component of the blockade occurs when a nucleic acid occupies the 5 nm long stem of the channel, and neglect contributions by the larger vestibule. The pore volume within the stem is 18 nm^3, and the average diameter is 2.0 nm, with a limiting aperture of 1.5 nm at the neck of the pore.

Three parameters provide information about the nature of the linear polymer passing through a nanopore. The first is blockade amplitude, which is best normalized and expressed as a fraction or percentage of the open channel current, I/I_o, where I is the blockade current and I_o is the open channel current. I/I_o has a characteristic value for many homopolymers of RNA and DNA, suggesting that

it will be an important analytical feature of nanopore technology.

Although several factors could conceivably contribute to blockades of ionic current, the simplest to test experimentally is that the fractional volume of a linear nucleic acid strand occupying a pore will reduce the number of ions available to carry current. The contribution of fractional volume of the molecule occupying the pore stem has been tested by investigating blockades produced by several polyanions that vary in molecular volume yet have approximately the same charge density of phosphate along the strand. These blockade amplitudes were then compared with the fractional volume of the pore stem occupied by the polymer, taking into account not only the molecular volume of the polymer but also water of hydration on the polymer and pore walls. The exact conformation of a single strand of DNA in the channel is unknown, but for the purposes of this calculation a conservative estimate is 0.34 nm per base, the repeat distance of bases in a double helix. About 15 nucleotides would then be in the pore at any given time during translocation. (The pore is defined here as the portion of the hemolysin channel that penetrates the lipid bilayer and does not include the larger volume of the vestibule and channel mouth.) Assuming that each monomer of the nucleic acid has four waters of hydration (60 total), and that a single layer of water is bound to the interior surface of the pore (320 total), approximately 70% of the available pore volume would be occupied by a single strand of oligo(dA).

Similar calculations were made for oligo(dC), for an abasic strand of nucleic acid, and for polyphosphate, both of which also produce measurable ionic current blockades [15]. The experimental values are in approximate agreement with the calculated values. From the results of this simple experiment, it can be concluded that blockade amplitude is largely a function of the fractional volume occupied by a linear polymer traversing a nanopore. The difference in total volume between purine and pyrimidine deoxyoligonucleotides in the α-hemolysin pore is only 0.3 nm^3, which represents a 6% difference in the volume occupied by the molecules in the pore after correcting for water of hydration. This difference is just barely detectable as an average signal over noise and is produced by multiple nucleotides occupying the length of the pore. It follows that an improved pore having a smaller limiting aperture than that of the α-hemolysin channel will probably be required for single nucleotide resolution in nanopore sequencing applications.

6. CORRELATION OF BLOCKADE PARAMETERS WITH POLYNUCLEOTIDE COMPOSITION

If blockades are caused by the transport of linear polynucleotides through the pore, their lifetimes should be proportional to the length of the polymer. This was tested for polyU, polyC, and polyA [10], and it was found that all three homopolymers exhibited a linear relation between chain length and duration of the blockade. Akeson et al. [16] went on to investigate the amplitude of blockades caused by homopolymers of RNA. The pattern of blockades caused by polyC was easily distinguished from the pattern for polyA, whether the polymers were examined separately or in a mixture. The channel current was reduced significantly more by polyC RNA (typically 95% blockades) than by polyA RNA (84% blockades), and the polyC blockades were shorter in duration, averaging 6 μs per nucleotide compared to 16 μs per nucleotide for polyA.

The fact that polyA and polyC could be distinguished by the nanopore suggested that a transition from polyA to polyC segments within a single RNA molecule should be detectable. An RNA was synthesized with the sequence $A_{(30)}C_{(70)}Gp$, which would be expected to produce a bilevel blockade. Such blockades were in fact commonly observed and interpreted to represent an initial blockade by polyC RNA (current reduced to 5% of open channel value) followed by polyA RNA (current reduced to 16% of open channel value).

The evidence presented by Akeson et al. [16] provided the first demonstration that modulations of ionic current during translocation of a macromolecule through a pore can be used to discriminate between components along its linear sequence. At first the greater blockade amplitude associated with polyC seemed paradoxical because adenine (a purine) is a larger base than cytosine (a pyrimidine). If the size of individual nucleotides within the RNA strands alone dictated blockade amplitudes of ionic current, the polyA blockade might be expected to have had a greater amplitude. The paradox was resolved when it was recalled that polyC forms a single-stranded helix with a 1.34 nm diameter [17] that can readily translocate through the pore as a helix. PolyA also forms a single-stranded helix, but with a diameter of 2.1 nm [18]. It therefore seems likely that polyA penetrates the pore only when the helical strand is partially unwound into an extended configuration, which also permits more ionic current to pass during a blockade. The additional energy required to unwind the purine bases is consistent with the significantly longer blockade duration associated with polyA strands.

7. NANOPORE ANALYSIS OF DNA AT SINGLE BASE-PAIR RESOLUTION

From the measured velocities of single stranded nucleic acid translocation through the hemolysin pore, it is clear that a major factor limiting resolution is the time available to detect current modulation by a single nucleotide. For instance, a single nucleotide in a DNA strand passes through the length of the pore in a few microseconds, a time interval that does not provide sufficient ionic current to distinguish between a purine base and a pyrimidine base in the total ionic current of a blockade. In order to demonstrate that increased time intervals may permit the sensitivity required for single base resolution, DNA "hairpins" can be used to keep a single molecule of DNA in the α-hemolysin channel for relatively long (millisecond) time intervals. Hairpins are sequences of single-stranded nucleic acids folded back on themselves to form hydrogen bonds between Watson-Crick base pairs [19, 20]. The entropy and enthalpy components of DNA hairpin formation are readily estimated, and the sequence can be designed so that only intramolecular interactions occur [21]. An extensively hydrogen bonded hairpin

structure can enter the vestibule of an α-hemolysin channel but would not be translocated through the pore until all the hydrogen bonds stabilizing the hairpin spontaneously dissociated.

The investigation reported by Vercoutere et al. [22] used a well-characterized DNA hairpin with a six-base-pair stem and a four-deoxythymidine loop. When captured within an α-hemolysin nanopore, this molecule caused a partial current blockade (or "shoulder") lasting hundreds of milliseconds, which was followed by a rapid downward spike. This signature was interpreted as an initial capture of a hairpin stem in the vestibule which produces the shoulder, followed by simultaneous dissociation of the six base pairs in the hairpin stem. The extended single strand can then traverse the pore and produce the spike. With this assumption, a series of blunt-ended DNA hairpins with stems that ranged in length from 3 to 8 base-pairs (bp) was found. It was found that each base pair addition produced a marked increase in median blockade shoulder lifetime, and a downward trend in shoulder current amplitude, from I/I_o equal to 68% for a 3 bp stem to I/I_o equal to 32% for a 9 bp stem (Fig. 2). These results demonstrate that single base pair resolution can be achieved by monitoring the duration and reduction of ionic current blockades as the hairpin stem extends further into the vestibule with each additional base pair.

8. CORRELATION OF BLOCKADE DURATION WITH FREE ENERGY OF DUPLEX STABILIZATION

In the experiments described, the standard free energy of hairpin formation was calculated and compared with the median duration of hairpin shoulder blockades [22]. The median blockade duration correlated remarkably well with the free energy of formation. This correlation would not be expected if the force generated by the electric field was simply unzipping the hydrogen bonded base pairs of the hairpin and drawing the molecule through the pore. Instead the hairpin remains in the vestibule until all of the hydrogen bonded complementary bases dissociate simultaneously, at which point the unwound DNA strand is drawn through the pore to produce the spike. The time required before the hairpin dissociates (the shoulder) is an exponential function of the number of base pairs in the hairpin structure and is directly related to the calculated free energy that stabilizes the hairpin. Note that this relation between free energy and blockade duration also extends to a single nucleotide mismatch in the hairpin, in which the AT pair is replaced by an AA pair. This reduces the median blockade duration from $\sim$1 second to 10 milliseconds, showing that the nanopore can distinguish two DNA molecules that differ by a single nucleotide.

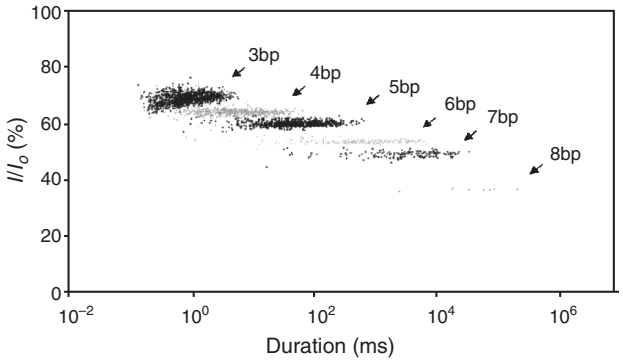


Figure 2. Nanopore resolution of single base-pair differences, illustrated as an event diagram for DNA hairpins with 3 to 8 base-pair stems. Each point represents the duration and amplitude of a shoulder blockade caused by one DNA hairpin captured in the pore vestibule. The duration of the 9 bp hairpin blockade shoulders lasted so long that a statistically significant number of events could not be readily obtained. Modified with permission from [22], W. Vercoutere et al., *Nat. Biotechnol.* 19, 248 (2001). © 2001, Nature.

These results illustrate the extraordinary ability of a nanopore to provide information about dynamic events occurring in single molecules. Furthermore, they demonstrate that ionic current blockades are sufficient to detect differences between DNA molecules at the single base-pair level, suggesting that a suitable nanopore may be able to resolve individual nucleotides.

9. NANOPORE ANALYSIS USING SEQUENCE-SPECIFIC DETECTION

Howorka et al. [23] found that it was possible to attach a single-stranded oligomer of DNA to hemolysin so that the vestibule contained a short sequence that potentially could bind to traversing DNA if complementary base pairing was possible. This remarkable feat of nanoengineering produced a biosensor that could detect specific DNA sequences with single base resolution. The principle underlying the function of this biosensor is that a traversing DNA strand binds to the tethered DNA in the vestibule by complementary base pairing and remains in the pore for a measurable amount of time. As a result, a distinctive two-part signal is generated, first by the duplex DNA that partially blocks the current, followed by a spike as the target DNA is released and passes through the pore stem. DNA that lacks the complementary base pairing passes through very rapidly without the partial blockade. The duplex lifetimes allow discrimination between short DNA strands that differ by only a single nucleotide so that duplex formation is inhibited. Because of this sensitivity, it was possible to determine the sequence of a codon embedded in the sequence of a single DNA oligomer covalently linked to the vestibule. These authors also used this approach to detect a mutation in the reverse transcriptase gene of the HIV virus.

10. COMPUTATIONAL APPROACHES TO ANALYSIS OF NANOPORE DATA

A key strength of nanopore detectors is that they not only detect and identify single molecules but can also analyze populations of single molecules. With signal processing and pattern recognition, this information provides an analytical tool based on a rapid sampling of channel current blockades. In recent work, a computational tool incorporating machine-learning methods has been used to perform signal analysis

and pattern recognition [22, 24]. Feature extraction methods, such as hidden Markov models (HMMs), offer a powerful and generic data-mining tool. For instance, labeling the HMM analyzed data-events, followed by support vector machine (SVM) discrimination, allows pattern recognition among thousands of individual blockade events. SVM is in many ways superior to neural nets, in that performance optimization is done with structural risk minimization (SRM). Due to their SRM constraint, SVMs are much less susceptible to overtraining, which permits a training process that is easily deployed and scalable.

Two hurdles must be overcome for successful signal processing by HMM and SVM approaches. At the earliest stage of the signal acquisition, ad hoc methods generally must be used. In recent channel current pattern recognition work, signal acquisition and primitive feature selection were performed by a time-domain finite state automaton (FSA) and a wavelet-domain FSA. The second hurdle is the data labeling needed for training the discriminator (supervised learning). In the case of nanopore experiments the labeling problem was solved by using training data drawn from operation with only a single type of analyte, so that all training events had the same label.

A primary computational goal of the blockade pattern recognition was real-time signal calling, in which event patterns are identified in less time than the duration of the measurement. This goal has practical use in detector operation in that extensive data capture is not required. In other words, detector data outflow does not exceed the throughput of the signal processing pipeline. In recent work involving 100 ms blockade signatures, it has been demonstrated that real-time signal calling is possible with an inexpensive PC using Linux (a free Unix-type operating system), and the C and Perl software packages.

11. SYNTHETIC NANOPORES FOR ANALYZING NUCLEIC ACIDS

Synthetic nanopores are now being tested that could have the potential to match or surpass α-hemolysin in detecting and analyzing nucleic acids in solution. In contrast to a labile protein in a delicate lipid bilayer, solid state nanopores should be mechanically robust, tolerate a broad range of temperatures, pH, and chemical conditions, and provide a low capacitance, low noise surface suitable for integrated electronics. Ideally, the fabrication process used to make the solid state nanopore should make it possible to easily control and vary pore dimensions so that these dimensions can be optimized for a particular application (e.g., probing single-stranded vs. double-stranded DNA).

Fulfilling these requirements is difficult because fabrication methods able to manipulate matter at nanometer dimensions have not been available. For example, single 3–5 nm holes in an insulating silicon nitride membrane can be made using an erosion process, such as reactive ion etching [25], but this method produced nanopores with a range of dimensions from 3 to 40 nm, only a few of which were of the desired dimension. Achieving greater reproducibility requires knowing precisely when to stop the erosion process. One approach that can reproducibly fabricate nanopores of a desired dimension is a feedback-controlled ion sputtering system that counts the ions transmitted through the gradually opening pore and extinguishes the ion sputtering erosion process at the appropriate time [26]. With feedback control, reproducibility does not depend on precisely matching all conditions and starting dimensions. Using this system, it is possible to routinely produce robust single nanopores with 2, 3, or 4 nm diameters. Such nanopores are capable of registering single DNA molecules that produce characteristic current blockade signals in much the same way as they do in the α-hemolysin pore [23].

12. FUTURE APPLICATIONS OF NANOPORE ANALYSIS

To summarize, single-base resolution in DNA or RNA remains the goal of investigators now working to establish nanopore sequencing, and several significant intermediate applications are now within reach. These include simple detection of DNA in microscopic volumes, as well as targeted molecular barcodes. Although nanopore analysis is just emerging as a new research tool, the rapidly developing technology has immense promise as an analytical method for measuring and characterizing linear polyions such as nucleic acids. The ability to detect and identify single nucleic acid molecules in solution has numerous applications, but these will be realized only if a stable inorganic nanopore can be developed. We are making significant progress toward this end, and as noted the first such nanopore based on silicon nitride has detected long double-stranded DNA molecules as they pass through. Assuming that a nanopore instrument can be commercially developed, the following applications appear to be feasible:

1. Measurement of nucleic acid concentration, for instance, during PCR amplification.
2. Identifying molecular species in solution. A nanopore in the size range of 4–6 nm diameter will be able to detect and perhaps identify individual soluble protein molecules.
3. Encoded polymers as molecular markers. Akeson et al. (unpublished results) have demonstrated that synthetic nucleic acids can be linked to a cleavable targeting agent and used to identify and quantify analytes in solution. After binding of the targeted marker to the analyte, the linker is cleaved and the nucleic acid marker is determined by its nanopore signal.
4. Hybridized DNA is readily distinguished from single stranded DNA by a nanopore of appropriate size.

GLOSSARY

Alpha hemolysin A toxin produced by Staphylococcus bacteria. Hemolysin assembles into heptameric structures in lipid bilayers and produces a nanopore approximately 2 nm in diameter.

Base pair In the double helical structure of DNA and RNA, the complementary bases hydrogen bond to from base pairs. A bonds to T (DNA) or U (RNA) and G bonds to C.

DNA hairpins Short strands of DNA ~ 10–50 bases in length with complementary base pair sequences at both ends are able to assemble into a hairpin-like loop. The ends of the loop form the double helix structure of DNA.

Hybridization When two strands of DNA with complementary base sequences come together in solution, they form a double helix. This process is called hybridization.

Lipid bilayer Amphiphilic molecules such as phospholipids self-assemble into biomolecular structures called lipid bilayers. Such structures represent the permeability barrier of all cell membranes, and can be used to support a nanopore such as hemolysin.

Nanopore A nanoscopic pore (1–100 nm diameter) in a thin membrane.

polyA, polyC, polyU When RNA is synthesized from a single species of nucleotide, it is called a homopolymer. PolyA, polyC, and polyU are abbreviations from polyadenylic acid, polycytidylic acid, and polyuridylic acid, all RNA homopolymers.

Polymerase chain reaction (PCR) A method for enzymatically amplifying DNA.

ACKNOWLEDGMENTS

Portions of this chapter were adapted from recent reviews by Deamer and Branton [27] and Deamer et al. [15].

REFERENCES

1. W. A. Coulter, U.S. Patent 2, 656, 508, 1953.
2. R. W. DeBlois and C. P. Bean, *Rev. Sci. Instrum* 41, 909 (1970).
3. C. P. Bean, *Membranes* 1, 1 (1972).
4. R. W. DeBlois, C. P. Bean, and R. K. A. Wesley, *J. Coll. Interface Sci.* 61, 323 (1977).
5. S. B. Hladkey and D. W. Haydon, *Biochim. Biophys. Acta* 274, 294 (1972).
6. B. A. Cornell, V. A. Braach-Maksvytis, L. G. King, P. D. Osman, B. Raguse, L. Wieczorek, and R. J. Pace, *Nature* 387, 580 (1997).
7. S. M. Bezrukov, I. Vodyanoy, and V. A. Parsegian, *Nature* 370, 279 (1994).
8. S. M. Bezrukov, *J. Membr. Biol.* 174, 1 (2000).
9. B. Walker, J. Kasianowicz, M. Krishnasastry, and H. Bayley, *Protein Eng.* 7, 655 (1994).
10. J. Kasianowicz, E. Brandin, D. Branton, and D. W. Deamer, *Proc. Natl. Acad. Sci. USA* 93, 13770 (1996).
11. A. Meller, L. Nivon, and D. Branton, *Phys. Rev. Lett.* 86, 3435 (2001).
12. A. Meller, L. Nivon, E. Brandin, J. Golovchenko, and D. Branton, *Proc. Natl. Acad. Sci. USA* 97, 1079 (2000).
13. M. Muthukumar, *J. Chem. Phys.* 111, 10371 (1999).
14. D. K. Lubensky and D. R. Nelson, *Biophys. J.* 77, 1824 (1999).
15. D. Deamer, H. Olsen, M. A. Akeson, and J. Kasianowicz, "Structure and Dynamics of Confined Polymers," NATO Science Series, Vol. 87, p. 165. Kluwer Academic, Dordrecht, 2002.
16. M. Akeson, D. Branton, J. J. Kasianowicz, E. Brandin, and D. W. Deamer, *Biophys. J.* 77, 3227 (1999).
17. C. R. Cantor and P. R. Schimmel, "Biophysical Chemistry, Part III: The Behavior of Biological Macromolecules," p. 1371. Freeman, San Francisco, 1980.
18. W. Saenger, J. Riecke, and D. Suck, *J. Mol. Biol.* 93, 529 (1975)
19. D. A. Erie, A. K. Suri, K. J. Breslauer, R. A. Jones, and W. K. Olson, *Biochemistry* 32, 436 (1993).
20. V. P. Antao and I. Tinoco, Jr., *Nucl. Acids Res.* 20, 819 (1992).
21. D. Rentzeperis, K. Alessi, and L. A. Marky, *Nucl. Acids Res.* 21, 2683 (1993).
22. W. Vercoutere, S. Winters-Hilt, H. Olsen, D. Deamer, D. Haussler, and M. Akeson, *Nat. Biotechnol.* 19, 248 (2001).
23. S. Howorka, S. Cheley, and H. Bayley, *Nat. Biotechnol.* 19, 636 (2001).
24. S. Winters-Hilt, W. Vercoutere, V. S. DeGuzman, D. Deamer, M. Akeson, and D. Haussler, *Biophys. J.*, in press.
25. D. C. Ralph, C. T. Black, and M. Tinkham, *Phys. Rev. Lett.* 74, 3241 (1995).
26. J. Li, D. Stein, C. McMullan, D. Branton, M. J. Aziz, and J. A. Golovchenko, *Nature* 412, 166 (2001).
27. D. Deamer and D. Branton, *Acc. Chem. Res.* 35, 817 (2002).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanoporous Carbons

Yoshio Yamada
University of Fukui, Fukui, Japan
Jun-ichi Ozaki
Gunma University, Kiryu, Gunma, Japan

CONTENTS

1. INTRODUCTION

Carbon material is a versatile substance because of its variety of bonding character. The typical allotropic forms are graphite and diamond, in which carbon element consists of sp^2 and sp^3 hybridization, respectively. In general, carbonaceous products categorized in the former carbon are obtained when many organic compounds and those of biological origin such as wood, peat, coal, and biomass are pyrolyzed in inert atmosphere at high temperature. The products contain a lot of elemental carbon and a small amount of oxygen, nitrogen, and so on.

Franklin classified the carbonaceous materials into two distinct groups (e.g., nongraphitizable and graphitizable carbon) on the basis of X-ray diffraction study [1]. Also hard and soft carbons are often referred to under another nomenclature, respectively [2]. The structural models proposed are represented schematically in Figure 1. Thereafter the great advance of the high resolution transmission electron microscope (TEM) has made it possible to visualize the more precise structure of the carbons. Jenkins and Kawamura observed TEM images of a glasslike carbon and illustrated a schematic model for the network of ribbon stacks, as shown in Figure 2a [3]. In addition Shiraishi proposed another model for nongraphitizable carbon (Fig. 2b) [4]. Figure 3 shows typical lattice images of nongraphitizable carbon from phenol formaldehyde resin and of graphitizable carbon from petroleum pitch coke. They reflect the skeleton of the hexagonal aromatic carbon layers [4]. The resin carbon does not form a structure like Figure 3d even at high temperatures like 3273 K, whereas the number of parallel layers increases considerably in the pitch carbon with heat treatment at much lower temperature, indicating a better graphitic structure.

It is said that such a structural difference originates both from the structure of the raw material and from the carbonization process: the cross-linking develops between the randomly oriented crystallites during carbonization of the resin leading to a rigid structure and to a charcoal which is hard carbon and has a well developed porous structure. On the other hand, the pitch proceeds in a liquid state during the carbonization around 627–773 K and aromatic hydrocarbons produced at this temperature range can arrange in parallel with each other to lead a formation of small spherical droplets with optical anisotropy called carbonaceous mesophase [5]. It has been shown that oxygen-rich precursors inhibit graphitization whereas high C–H ratios enhance the formation of a three-dimensional graphitic structure with less developed porous structure. The nongraphitizing carbon contains a large number of pores between small crystallites as can be seen from Figure 1a. Most of these pores originally included are located inside far from the surface. Thus, in order to generate the open pores on the surface of the particles, carbon atoms are partially burned off at high temperature using oxidative agents such as steam, carbon dioxide, and so on. The process and the resultant product are referred to as activation and activated carbon. Since most of the pores are mainly the spaces between a few stacked lamellae or crystallites, the pore shape in the carbon materials is basically of slitlike structure.

Activated carbon is one of the well-known porous substances and a collective name for a group of porous carbons. In 1911, Norit Company began the commercial production

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (237–262)

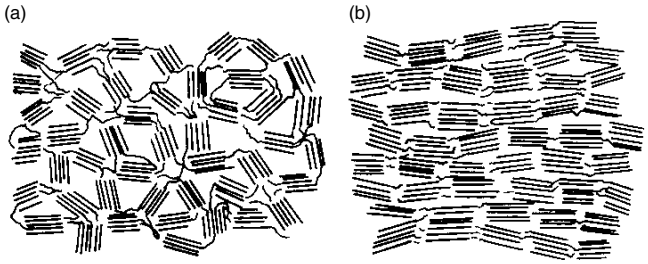


Figure 1. Schematic representation of (a) nongraphitizable and (b) graphitizable carbons. Reprinted with permission from [1], R. E. Franklin, *Proc. Roy. Soc. London Ser. A* 209, 196 (1951). © 1951, The Royal Society of London.

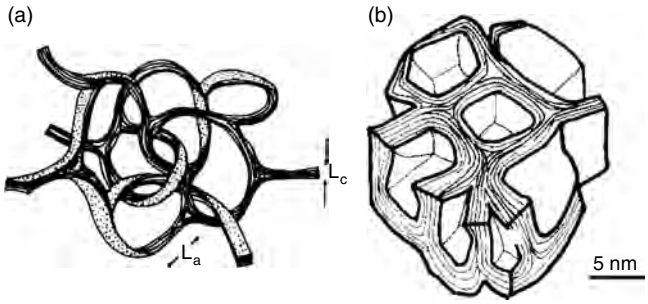


Figure 2. (a) Structural model for the network of ribbon stacks in glassy carbon by Jenkins and Kawamura, L_a: lattice constant of *a*-axis, L_c: lattice constant of *c*-axis. Reprinted with permission from [3], G. M. Jenkins and K. Kawamura, *Nature* 232, 175 (1971). © 1971, Nature. (b) Structural model for nongraphitizable carbon by Shiraishi. Reprinted with permission from [4], M. Shiraishi, in "Introduction to Carbon Materials" (Carbon Society of Japan, Ed.), revised version, p. 29. Science and Technology Publishing, Tokyo, 1984 [in Japanese]. © 1984, The Carbon Society of Japan.

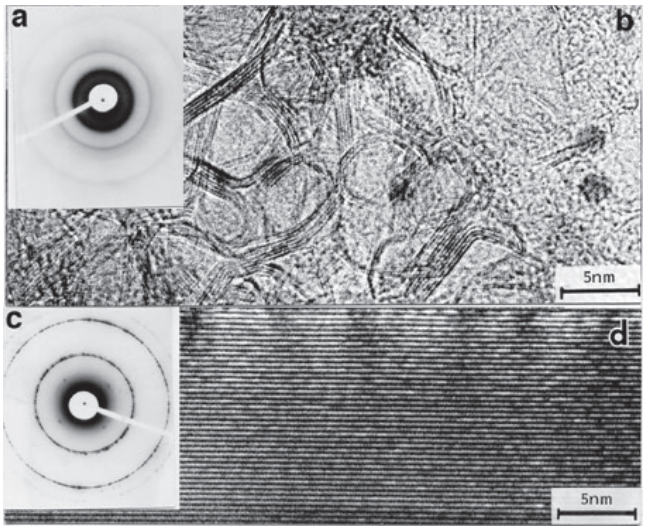


Figure 3. Electron microdiffraction (a) and their TEM lattice image (b) of nongraphitizable carbon from phenol formaldehyde resin, and those (c and d) of graphitizable carbon from petroleum coke. Reprinted with permission from [4], M. Shiraishi, in "Introduction to Carbon Materials" (Carbon Society of Japan, Ed.), revised version, p. 29. Science and Technology Publishing, Tokyo, 1984 [in Japanese]. © 1984, The Carbon Society of Japan.

of activated carbons from peat by steam activation. During the first decades of the last century, activated carbon was used mainly for the purification of products of the chemical, pharmaceutical, and food industries. Purification of drinking water was also an important application from the outset. In recent years, the production of carbon has been increasing for the prevention of environmental pollution and for meeting the constantly increasing demands for purity of natural and synthetic products [6].

Porous carbon materials have so far played an important role in many industries, including air and water purification, gas separation, catalysis, chromatography, and so on. However, there are numerous other potential applications in which materials with carbonaceous surfaces would be attractive, for instance adsorption of large hydrophobic molecules such as vita dyes, humic acids, dextrines, chromatographic separation, electrochemical double layer capacitors (EDLC), and lithium batteries. In these cases, the presence of wider pores, preferably mesopores, would be more advantageous to the adsorption of these large compounds.

To meet such requirements many novel approaches have been proposed to control pore structure and to produce different types of porous carbons. And it has been demonstrated that such porous carbons are very useful and exhibit outstanding performance in many new applications such as catalytic supports [7] and EDLC with very high power [8].

In the first section of this chapter the general definition of a pore in solid materials and the characteristics of the pores included in carbon materials are briefly represented to help readers understand the contents described in the following sections. Second, the importance of the nanopores in the carbons is pointed out and then special emphasis has been put on the novel preparations of nanoporous carbons and development of new utilization areas of these carbons. Throughout the chapter readers can survey the state of the art preparation and application for nanoporous carbons.

From this point of view carbon nanotubes and nanohorns are also promising materials because the pores have a unique shape and are more uniform than those in the conventional carbons.

Here the authors point out the importance of mesopores (or nanopores) in the porous carbons. All the micropore entrances in fibrous activated carbon [or activated carbon fiber (ACF)] can access the outer surface of the fiber, whereas most of the micropores existing in granular activated carbons are placed deep inside far from the surface of the carbon particles through the mesopores. Such a difference in the circumstance of the pores in fibrous and granular carbons is depicted in Figure 4. Of course, micropores are important to the adsorption ability of small molecules because most of the adsorption takes place in the micropores. In the case of granular carbons the species to be adsorbed have to approach to the micropores through meso- or macropores. In other words, these relatively large pores can serve as a passage for the transport of the adsorbates. In particular the presence and the quantity of the mesopore are very important in designing nanoporous carbons suitable for heterogeneous catalysis and supercapacitor applications.

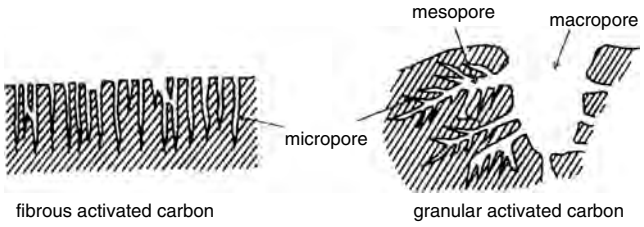


Figure 4. Structural models of pores in fibrous activated carbon and granular activated carbon.

2. CHARACTERIZATION OF NANOPOROUS CARBONS

2.1. Pore Definition

The general classification of pore in adsorbents is given according to the International Union of Pure and Applied Chemistry recommendation, dividing the size into the following groups (Table 1) [9];

- micropores: pore size <2.0 nm
- mesopores: pore size in the range of 2.0–50 nm
- macropores: pore size >50 nm

2.1.1. Micropores

A more precise classification of the micropores would distinguish two types of micropores, narrow and wide pores [10]:

- ultramicropores: pore size <0.7 nm
- supermicropores: pore size in the range of 0.7–2.0 nm

Micropores in porous carbons are very important for gas adsorption because their quantity and size are closely related to the specific surface area and the interaction with adsorbates. As a consequence of the overlapping of the adsorption potential of opposite pore walls, the filling of the microporosity corresponds to the initial steep portion of the adsorption isotherm at low relative pressure. The classification of adsorption isotherm curves will be explained in detail in Section 2.3.1.

Table 1. Classification of pore.

1. Classification by size (see text)
   - Micropore <2.0 nm
     Ultramicropore <0.7 nm
     Supermicropore 0.7–2.0 nm
   - Mesopore 2.0–50 nm
   - Macropore >50 nm
2. Classification by origin
   - Intraparticle pore: existing in crystallites
     Intrinsic pore: originated from crystal structure
     Extrinsic pore: induced by doping of foreign substances
   - Interparticle pore: generating by aggregation of primary particles
     Rigid pore
     Flexible pore
3. Classification by pore situation
   - Open pore: accessible to adsorbates
   - Latent pore: open, but very narrow so as to be inaccessible to adsorbates
   - Closed pore: isolated from surface and inaccessible to adsorbates

Source: Reprinted with permission from [16], M. Inagaki and K. Kaneko, in "Commentary on Carbon Family—Variety and Evaluation" (M. Inagaki, Ed.), p. 156. Agune Syoufuu Sha, Tokyo, 2001. © 2001, Agune Syoufuu Sha Publishing.

2.1.2. Mesopores

These pores with widths from 2.0 to 50 nm were previously named transitional pores, in which an adsorbate condenses in liquidlike state by capillary condensation and a meniscus is formed; as a consequence, a hysteresis loop often appears on desorption and its interpretation can lead to the distribution of mesopores in the carbon.

2.1.3. Macropores

Since these pores have widths exceeding 50 nm, they cannot be filled by capillary condensation. The limit of 50 nm is artificial and it corresponds to the practical limit of the method for the pore size determination based on the analysis of the hysteresis loop. These pores are not so important for the adsorption of gas and small molecules, but recently Toyoda and Inagaki have found that the macropores in exfoliated graphite have a high capacity for sorption of heavy oil [11]. In general, however, these large pores cause a decrease of the bulk density, one of drawbacks for transport usage such as natural gas storage vehicles [12].

The "nanopore" used in this chapter is not a formal nomenclature, but recently this word appeared in some papers as the importance of nanopores in the carbons is recognized [13–15]. The authors define it as pores covering the size dimension with a single nanometer ranging from larger micropores (1.0–2.0 nm) to smaller mesopores (2.0–10 nm). However, the size is not a restricted definition but has a certain range. Thus, in some cases we would include pores ranging from micropores to mesopores in the category of nanopores.

Pores in carbon can be further divided into the following two categories from a different point of view. The pore classification reported by Inagaki and Kaneko [16] is partially modified and shown in Table 1, together with the definition of pore size mentioned previously. The first category is separated into two groups (i.e., intra- and interparticle pore) by origin of the pore formation. The former, intraparticle pore, being present between particles (or crystallites), is further subdivided into two types of pore, intrinsic and extrinsic. The intrinsic pore originated from imperfection of the crystal structure itself, and the extrinsic one is induced by doping a foreign element such as boron or nitrogen. A typical example is the very narrow pore formed when a few layers stacked in graphite are expanded by intercalation. The interparticle pore can be classified into two pores due to the strength of the pores, hard and flexible pores. Most of the pores in the carbonaceous material (i.e., those in carbon aerogel) are hard because they are heat-treated at high temperature. But the pores formed by intercalation or those included in some original coals are considerably soft and easily deformed due to shrinkage at low temperature or they are destroyed by external forces. So when materials with such soft pores are employed for liquid nitrogen adsorption or mercury porosimetry experiments one should be careful of handling or data treatment.

In the second category pores are divided into open and closed pore. Pores measured by gas adsorption experiments are usually classified as open pores, but all of the pores by adsorption are not open, depending on the size of the probe molecule used. Very narrow pores, which the adsorbate molecules cannot access, fall into latent pores. Yoshizawa et al. have investigated accessible and inaccessible pores in microporous activated carbons by N_2 adsorption, density, and small angle X-ray scattering measurements [17]. The closed pores can be evaluated by TEM observation and small angle X-ray scattering which are described in the next section.

2.2. Characteristics of the Pores in Carbon Materials

As already mentioned in the Introduction, the small pores in carbon materials are spaces between stacked aromatic layers or interfacials/interstitials of the crystallites consisting of these stacked layers. Thus, it has been generally accepted that micropores in the carbonaceous materials like activated carbons possess slitlike structure (Fig. 1a).

The carbon nanotube was discovered by Iijima in 1991 [18]. Since a single walled carbon nanotube (SWNT) is composed of a monolayer graphene sheet rolled up into a long seamless cylinder, the substance has a structure of packed bundles of the tubes, in which the pores are located inside of the tubes as well as at intertubular channels. These pores shown in Figure 5 are of round shape, being considerably different from those in the conventional carbons [19]. In 1999 also Iijima et al. have found an analogous type of nanotube called a nanohorn, which was obtained by CO_2 laser ablation of a graphite rod [20]. TEM images of particles prepared suggest that each particle consists of a tubularlike structure with cone caps. Based on the TEM observation they proposed a model for the packing of carbon nanohorns in a dahlia particle [20]. Furthermore, Murata et al., have measured the N_2 adsorption isotherms of these nanohorns and reported geometrical models for the horn and its aggregate structure shown in Figure 6 [21]. Accordingly, these modern carbon materials have round-shaped pores, as easily supposed from the structure of Figures 5 and 6.

2.3. Analysis Method and Limitation

The major techniques to characterize pores in carbon materials are indicated in Figure 7, in which arrows give how much range the individual analytical technique can cover in the pore size included in the carbon. Each technique has various advantages and disadvantages as well as the limitation. A brief explanation and the limitation for each method are described.

2.3.1. Gas Adsorption

Gas adsorption is widely used as a standard method to analyze the porosity in carbon; the nitrogen adsorption/desorption isotherms are measured at the boiling point of liquid nitrogen. In recent years since almost all commercially available apparatus are equipped with computer-aided controller systems, one can collect isotherm data automatically. At the same time, pore characteristic parameters such as Brunauer, Emmett, and Teller (BET) surface area, pore

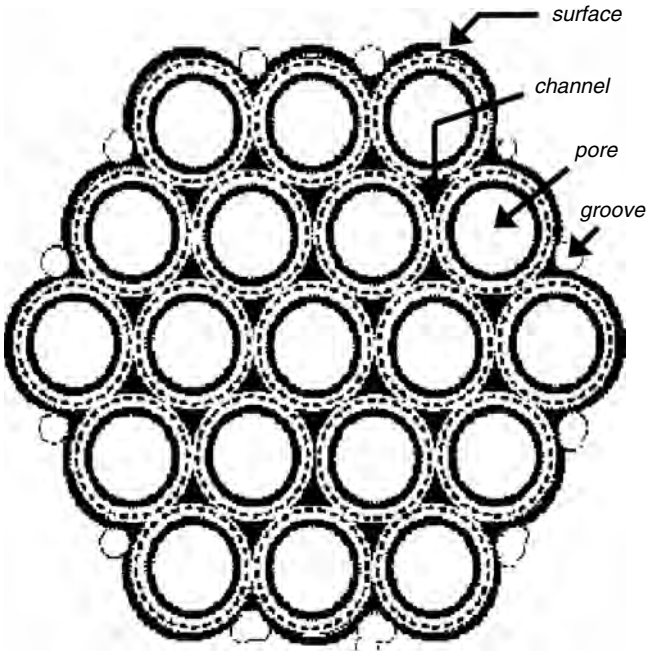


Figure 5. Schematic structure of a single nanotube bundle showing the pore, groove, channel, and surface sites. Reprinted with permission from [19], B. K. Pradhan et al., *Mater. Res. Soc. Symp. Proc.* 633 (2001). © 2001, Materials Research Society.

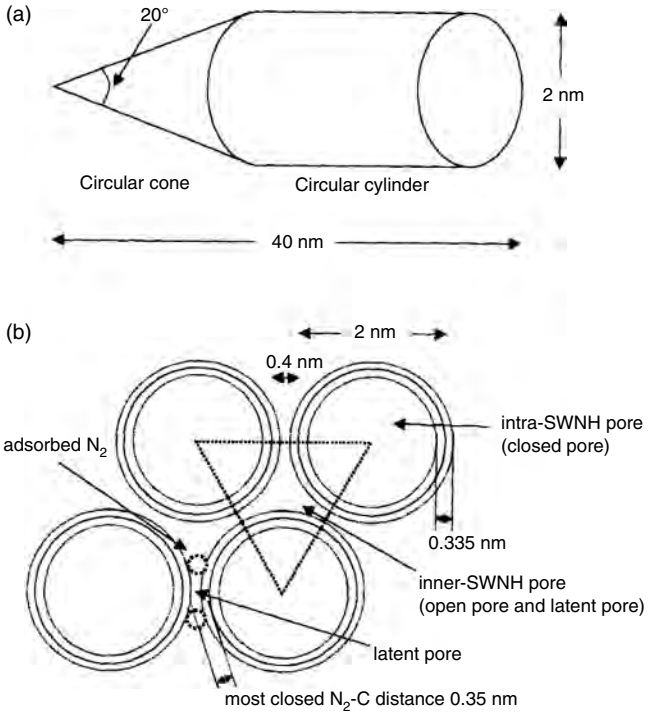


Figure 6. Geometrical models. (a) The model of the horn. (b) The model of the aggregate structure. Reprinted with permission from [21], K. Murata et al., *Chem. Phys. Lett.* 331, 14 (2000). © 2000, Elsevier.

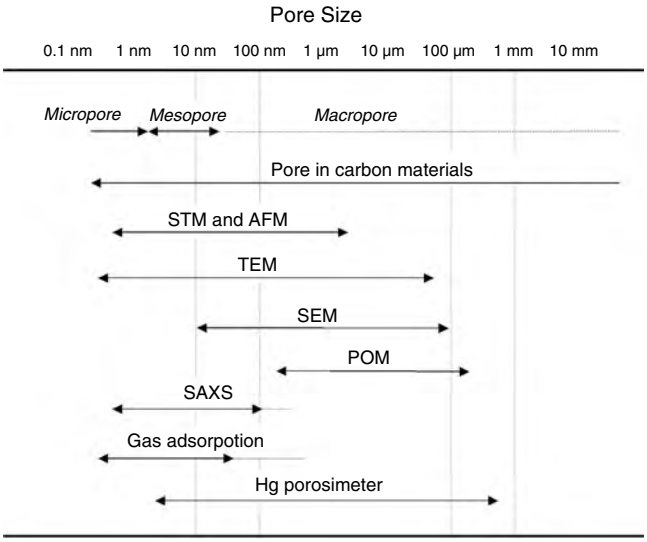


Figure 7. Range capable of evaluating pore size by various techniques. The detail of these techniques is described in each section. Reprinted with permission from [16], M. Inagaki and K. Kaneko, in "Commentary on Carbon Family—Variety and Evaluation" (M. Inagaki, Ed.), p. 156. Agune Syoufuu Sha, Tokyo, 2001. © 2001, Agune Syoufuu Sha Publishing.

size distribution, and so forth can be easily obtained by computer analyzing programs.

Isotherm Curve The shape of the adsorption isotherm curve reflects the characteristics of the pore structure in the adsorbents. Thus, it is very important to understand the relationship between the shape of the isotherm curves and pore structure. Brunauer et al. classified the curves into five groups (i.e., Types I–V) shown in Figure 8 [22].

Type I Isotherm The main feature of Type I is a sharp increase in adsorption at very low relative pressure (p/p_0) and a following long horizontal plateau over a wide range of higher pressure. This type is also called a Langmuir curve after his detailed analysis [23].

In this curve a monolayer adsorption takes place at very low pressure, which is indicative of strong interaction between the gases introduced and the micropore surface. The isotherm of the Type I can be found in porous carbons with a very narrow range of micropores (e.g., ACF, in which there is an enhancement of adsorption potential).

Type II Isotherm The isotherm of Type II is associated with monolayer–multilayer adsorption on an open and stable external surface of a powder, which may be nonporous, macroporous, or even, to a limited extent, microporous. The curve can be seen when an energetic heterogeneity in the adsorbent–adsorbate interactions can exist.

Type III Isotherm The isotherm of this type is confined to a few systems, in which the overall adsorbent–adsorbate interaction is weak, in contrast to relatively strong adsorbate–adsorbate interaction. For example, isotherm curves of this type can be observed when steam is adsorbed on the nonporous solids with a hydrophilic surface like graphite. The monolayer density tends to be unevenly distributed on the adsorbent surface with a relatively high concentration of

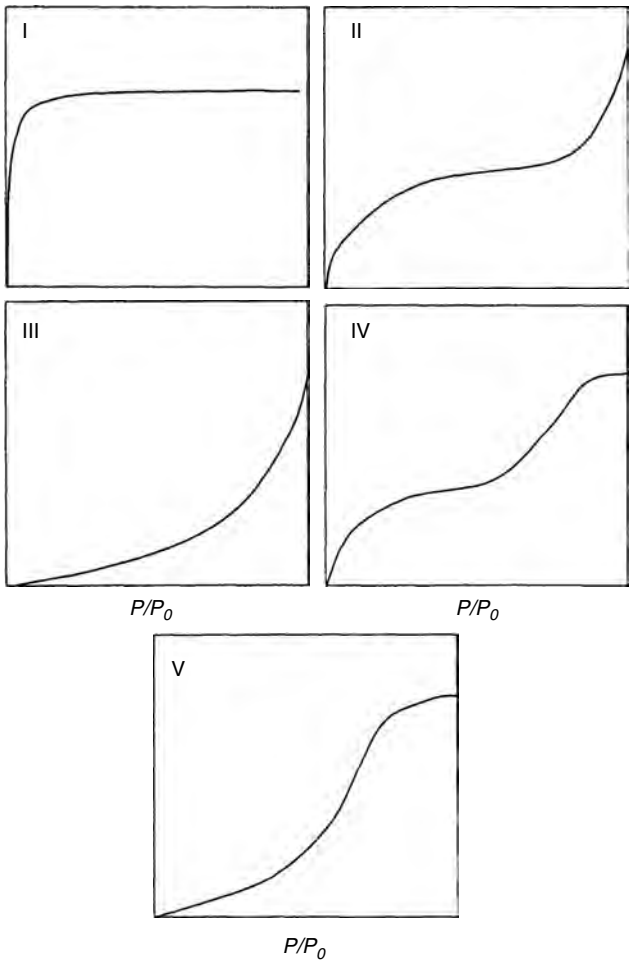


Figure 8. Shape of gas adsorption isotherms of Types I, II, III, IV, and V.

molecules located on the most active areas. As the pressure is raised the average monolayer concentration is increased, but before it can become closed-packed over the complete surface, the monolayer coverage is overtaken by a form of cooperative multilayer adsorption in which molecules are clustered around the most favorable sites.

Type IV Isotherm The isotherm curve can be observed in the case of porous solids with both mesopores and macropores, in which there are any interactions between adsorbent and adsorbate, being analogous to the behavior in the case of Type II. The typical example can be seen in the adsorption of nitrogen on activated carbons at 77 K. In general, the curves give a hysteresis loop between adsorption and desorption curves at relative pressure, $p/p_0 > 0.4$, which is due to the occurrence of capillary condensation at the adsorption into mesopores.

Type V Isotherm The appearance of the initial section of Type V isotherm is very similar to that of a Type III isotherm for a gas–solid system (e.g., water/carbon). In this case, however, the sharp increase in adsorption at higher p/p_0 is dependent on the pore size. For example, the ultramicropores in a molecular sieve carbon are filled with water at a much lower p/p_0 than are the wider pores in a supermicroporous carbon.

Surface Area In order to determine the surface area of porous carbons, the well-known BET method is always used as a standard procedure [24]. In principle, the BET method can be applied to isotherms of all types described previously, provided that there is no significant primary micropore filling contribution and that the structure of the completed monolayer does not vary from one surface to another.

Micropore Analysis Various procedures have been used to evaluate the micropore capacity for the experimental isotherm data, but in practice these are all empirical methods. Dubinin utilized the characteristic curve principle of the Polanyi potential theory [25, 26]. The fractional filling, n/n_0, of the micropore volume is expressed in the general form

$$n/n_0 = \exp -(A/E)^m$$

where m is a small integer, E is a characteristic free energy of adsorption for the given system, and A, originally termed adsorption potential, is the change in differential free energy of adsorption, viz.:

$$A = -RT \ln(p/p_0)^2$$

In this equation, R is a gas constant and in the special case of $m = 2$, the adsorption isotherm equation in linear form becomes

$$\log n = \log n_0 - D(\log p/p_0)^2$$

where D is a constant related to E. This equation is generally known as the Dubinin–Radushkevich (DR) equation. According to the DR equation, a linear relationship should be obtained between $\log n$ and $(\log p/p_0)^2$, with the intercept equal to $\log n_0$, the mciropore volume.

t-Plot and α_s Plot The t plot introduced by Lippens and de Boer [27] is one of the methods to compare experimental isotherms with the standard isotherm obtained on a suitable nonporous reference adsorbent. It provides a simple and direct method of comparing the shapes of the adsorption isotherms of the sample with that of the standard nonporous reference from which valuable information on the porous structure of the adsorbent can be obtained. The amount adsorbed on the porous adsorbent is plotted against t, the corresponding multilayer thickness of the standard nonporous solid. Actually the adsorption isotherms for a given adsorptive on different adsorbents may be superposed by plotting them in the reduced form, $t(n/n_m)$, against p/p_o. The backextrapolation to the ordinate of the linear portion of the plot gives a measure of the total micropore volume, whereas the slope of the linear section provides a measure of the more external area of the solid and the extrapolation to the origin of the initial part of the plot provides the total surface area.

The α_s method is another empirical approach for the determination of the microporosity [28]. The application of the method requires a standard isotherm of the adsorption (nitrogen in this case) on nonporous reference materials with a similar surface structure to the carbon under investigation. The amount adsorbed is plotted against the reduced standard adsorption (placing $\alpha_s = 1$ at $p/p_o = 0.4$).

In using both the t plots and α_s plots for the characterization of microporous carbons, one should understand that they are based on the assumption that adsorption in micropores is complete before it starts on the rest of the surface. Very often, however, depending on the pore structure of the carbon, there is a wide range of relative pressure where significant amounts of adsorption are occurring in both micropores and mesopores.

Mesopore Analysis The mesoporosity of the carbons can be evaluated by the analysis of the hysteresis loop observed on desorption, following the adsorption of vapors to initial relative pressure. The hysteresis loop results from the progressive condensation of the adsorbate in the pore, in the course of adsorption, with the formation of a meniscus. The vapor pressure above the meniscus decreases with the pore size, due to the small radius of curvature, and it follows that the smallest pores are desorbed last when the pressure in the vapor phase is reduced during desorption. The pore size distribution of mesopores can be calculated from the hysteresis loop by the computational techniques described in detail by Gregg and Sing [29]. The mesopore capacity is the amount adsorbed at the plateau of the Type IV isotherm (Fig. 8): to obtain the mesopore volume, it is assumed that the condensate has the same density as the liquid adsorptive at the operational temperature. There are various classical procedures for mesopore size analysis, such as the Barrett, Joyner, and Halenda method [30], the Cranston–Inkley method [31], and the Dollimore–Heal method [32, 33]. The details of these methods are given in each paper.

2.3.2. Analysis by Small-Angle X-ray Scattering

Since small-angle X-ray scattering (SAXS) occurs at the local electron density difference in materials, either pore or matrix can be scattered in principle. Therefore it provides averaged information over all pores regardless of their accessibility [34]. Recently, a pore characterization method which combines SAXS and N_2 adsorption with density data has been proposed and applied to pitch-based ACFs prepared with different activation extents [35]. The difference including inaccessible pores is successfully described among the ACFs by this method. In this section the SAXS results on the pore structure of carbon fibers prepared by a polymer blend are introduced as an example.

Fukuyama et al. applied the SAXS technique to evaluate the pore shape and size of the carbon fibers prepared from two kinds of polymers, in which one is phenolic resin as a carbon precursor, and the other is polystylene that disappears by pyrolysis, (e.g., a pore forming resin) [36]. Two kinds of samples heat-treated at 773 and 1273 K (sample A and B, respectively) were used for the measurements. The SAXS patterns are shown in Figure 9, where I and s are intensity in arbitrary units and a scattering parameter, respectively. In order to analyze the radius of gyration, R_g, in the pore by the Guinier and Fournet method [37], the regression lines to fit the experimental data of the lower s^2-region were drawn as shown in Figure 10. The values (R_g) of 7.1 nm for sample A and of 6.9 nm for sample B were obtained from the slop of the straight lines in the smaller s^2-region. The information on the shape of pores is included

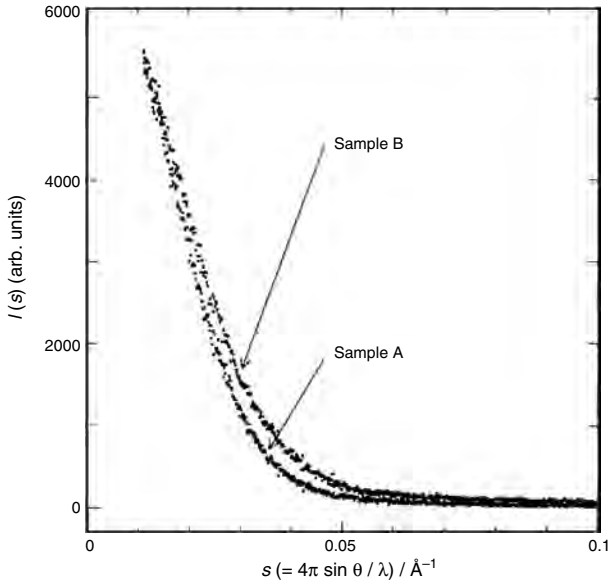


Figure 9. Scattering patterns of sample A (773 K) and sample B (1273 K). Reprinted with permission from [36], K. Fukuyama et al., *Carbon* 39, 287 (2001). © 2001, Elsevier.

in the scattering pattern in larger regions than the Guinier one [37]. Thus, they drew the cross-section plot given by the expression [38]

$$\ln(I(s) \cdot s) = \ln(I(0) \cdot s) - 1/2R_c^2 \cdot s^2$$

where R_c is a gyration radius of the bottom of the column. The R_c value of the pores from the slop if the plot of $\ln(I(s) \cdot s)$ vs s^2 forms a straight line (Fig. 11). From the

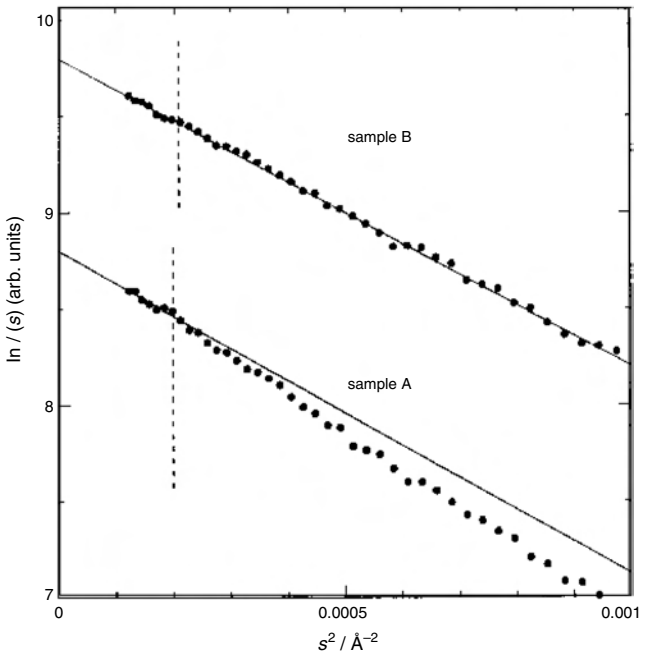


Figure 10. Guinier plots. The plot for sample B is displaced by +1 unit for clarity. Reprinted with permission from [36], K. Fukuyama et al., *Carbon* 39, 287 (2001). © 2001, Elsevier.

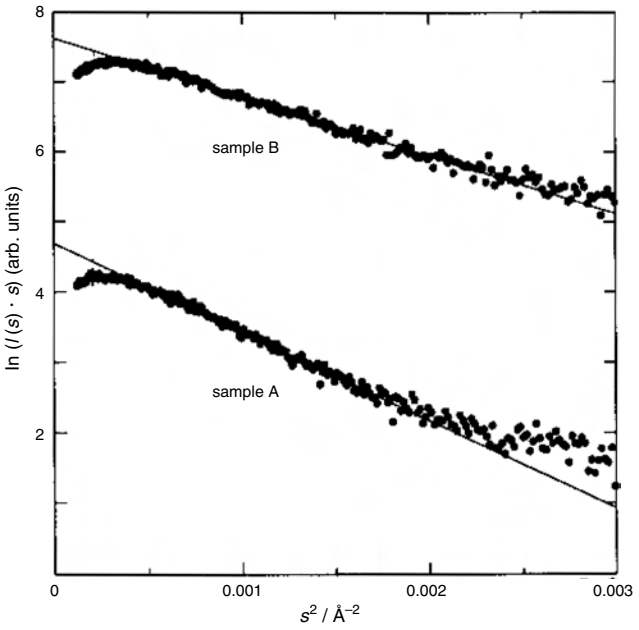


Figure 11. Cross-section plots. The plot for sample B is displaced by +3 unit for clarity. Reprinted with permission from [36], K. Fukuyama et al., *Carbon* 39, 287 (2001). © 2001, Elsevier.

following relation between R_c and radius of column bottom R, $R = \sqrt{2}\ R_c$, the value R can be calculated [38]. Also for the column with R and height L, the gyration radius R_c is given by

$$R_c^2 = (R^2/2) + (L^2/12)$$

Then, the values of L are determined as 17.4 and 19.2 nm for sample A and B, respectively. Table 2 and Figure 12 show the list of the results and the model of the pore structure in the carbon fiber heat-treated at 1273 K. Since the pore is elongated in the direction of the axis of a column and shrinks in the circle with heat treatment, the structural model is consistent with the results obtained by other methods.

2.3.3. Quantitative Analysis by TEM Observation

This method is useful for obtaining direct information on pore structure. Another advantage of this method is that a very small amount of the sample is enough for TEM observation. This tool has been used as a complementary technique for analyzing porous materials, because it gives only qualitative information on the pore character. However, recent advances in computer image analysis provide quantitative data on the pore structure.

Table 2. Structural parameters of sample A (773 K) and B (1273 K).

Sample	R_g^a (nm)	R_c (nm)	R (nm)	L (nm)	R_g^b (nm)
A	7.1 ± 0.3	5.0 ± 0.1	7.1 ± 0.2	17.4	7.2
B	6.9 ± 0.3	4.1 ± 0.1	5.8 ± 0.1	19.2	6.7

Source: Reprinted with permission from [36], K. Fukuyama et al., *Carbon* 39, 287 (2001). © 2001, Elsevier.

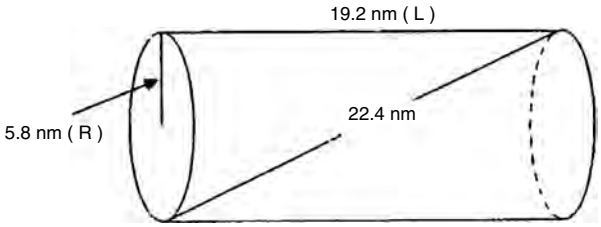


Figure 12. Structural model of pore in sample B (1273 K) based on the SAXS results. Reprinted with permission from the authors (Fukuyama et al.).

MacEnaney and Mays have applied quantitative microscopy or image analysis to the characterization of macropores by using as examples different types of porous carbon materials [39]. The principal advantage of the technique is an ability to evaluate both open and closed porosity. In addition, shape, location, and orientation of pores in the carbon matrix can be observed directly. It is also an advantage to be able to measure these parameters for different classes of macropores in a given porous body. The stereological problems of relating two-dimensional measurements to structural parameters of the three-dimensional material are a disadvantage, particularly when dealing with anisotropic and heterogeneous substances.

Here we introduce the analysis process of this technique and the results on the pore structure of ACF, which were carried out by Oshida et al. [40]. They used computer image analysis to improve the quantitative interpretation of the TEM images of ACF. Figure 13a shows the digitized picture of TEM images for the ACF sample obtained using an image scanner. The white parts correspond to the pores. The inset Figure 13b is the power spectrum of the fast Fourier transform of Figure 13a, and Figure 13c is the binary image of Figure 13(a) in order to clarify the pore images of the ACF. Finally the relationship of the pixel numbers in the outline (length of the peripheral of the binary images, X_N) against those inside the island (surface area of the pore, S_N) is shown in Figure 14. By using the following relation between X_N and S_N, $\log X_N = 1/2D \log S_N + A$, the fractal dimension (D) can be calculated to be $D = 1.74$, where A is a constant. On the basis of these results they proposed microscopic and macroscopic structural models for the pores in the ACF, as shown in Figure 15. The pores in the fiber connect together to form open pores, but the ultramicropores are isolated with each other. Consequently, this method is useful for the precise analysis of pore structure and will be more important with the advances of computer analysis.

2.3.4. Quantitative Analysis of STM and AFM Observation

Basically these methods are same as the analysis by TEM observation; the images obtained by scanning tunnel microscopy (STM) or atomic force microscopy observation (AFM) are treated statistically through computer analyses in order to evaluate the pore size distribution. Inagaki et al. applied the STM method to activated carbons from phenolformaldehyde resin particles and compared the pore size parameters obtained with those by the adsorption method [41–43].

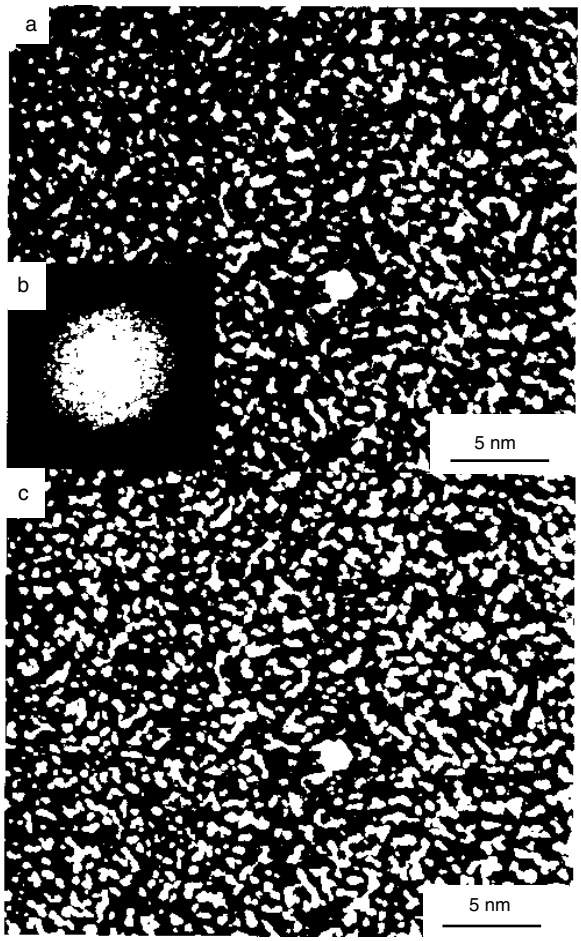


Figure 13. Digitized picture of a TEM image for ACF obtained using the image scanner. The inset (b) is the power spectrum of fast Fourier transform of (a), and (c) is the binary image of (a). Reprinted with permission from [40], K. Oshida et al., *J. Mater. Res.* 10, 2507 (1995). © 1995, Elsevier.

2.3.5. Quantitative Analysis of Scanning Electron and Polarizing Optical Microscopy Observation

The quantitative analysis of the pores is possible by scanning electron microcopy (SEM) or polarizing optical microscopy (POM) observation, but the magnification of these images is of μm order. Therefore these methods are effective for the macropore analysis, as is clear from Figure 7. For example, Inagaki and Suwa analyzed the macropore structure in exfoliated graphite by using SEM images [44]. Also the analysis of the pore in high density isotropic graphite was carried out by the POM method [45].

2.3.6. Mercury Porosimetry

Macroporosity and most mesoporosity can be assessed by mercury porosimetry, which is a simple technique based on the intrusion of mercury into the pores under the effect of pressure. The method can be used only when a sample to be measured has no wettability to liquid mercury. The range

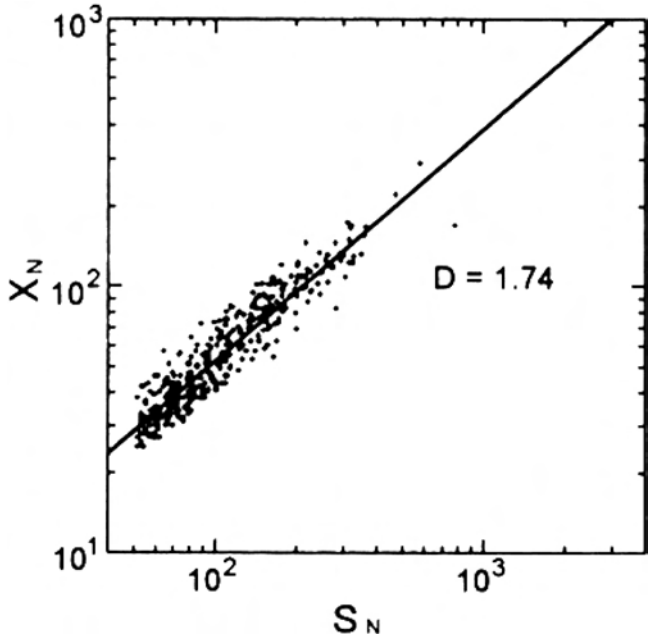


Figure 14. Plots of the pixel numbers in the outline of the binary images of ACF sample versus that inside the islands of the binary images (Figure 13c). Reprinted with permission from [40], K. Oshida et al., *J. Mater. Res.* 10, 2507 (1995). © 1995, Elsevier.

of mercury porosimetry extends from a pore size of about 7.5 nm at 200 MPa to 15 μm, the size of pores penetrated at atmospheric pressure. As pointed out in Section 2.1, when a soft sample with flexible pores such as a kind of raw coal is employed, one should be careful in the measurement.

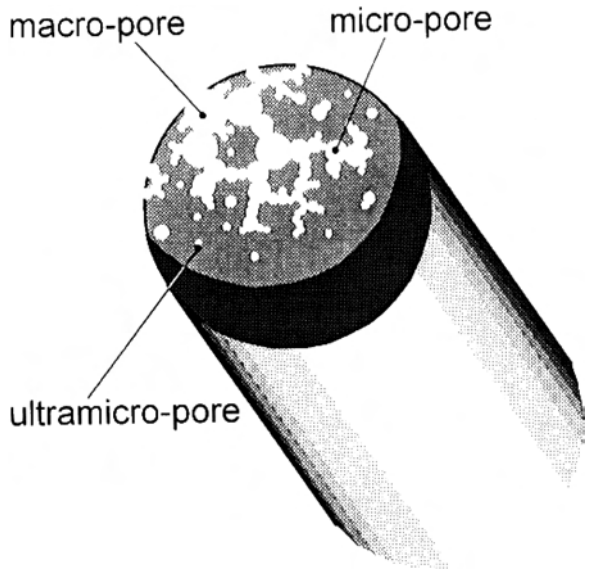


Figure 15. Macroscopic structural model of ACFs based on the presently visualized pore structure showing macropores, micropores, and ultramicropores formed by the activation process. Reprinted with permission from [40], K. Oshida et al., *J. Mater. Res.* 10, 2507 (1995). © 1995, Elsevier.

3. PREPARATION OF NANOPOROUS CARBONS

3.1. Conventional Preparation Methods

3.1.1. Carbonization

A brief description of conventional porous materials, activated carbons, is given in this section. Activated carbons are the carbon materials that possess well-developed pores in their bodies as shown in Figure 16 [46]. Carbon materials are usually prepared by heating organic materials, such as plant-based substance, synthetic polymers, etc., at around 1273 K in nonoxidative atmosphere. This heat treatment process is called carbonization, which is a collective name for the reactions like dehydration, cyclization, aromatization, condensation, and dehydrogenation, resulting in the formation of graphitic structure. For the preparation of porous carbon materials, activation treatment is needed.

Activation is a process to introduce pores into carbon materials. This is usually divided into two categories, physical activation and chemical activation [46].

3.1.2. Physical Activation

This process is a gasification of carbons by gaseous oxides. In other words, partial oxidation can leave holes or voids in the materials depending on the extent of the reaction. Steam and carbon dioxide are exclusively used for this purpose, because the basic reactions for these gases are endothermic which enables fine control of the gasification [47]. In this sense, oxygen is not a good gasifying reagent, because the reactions between carbon and oxygen have large exothermic heat of reaction. Steam can gasify carbons above 1023 K, while carbon dioxide can activate carbons above 1123 K, reflecting the difference in the specific rate of reaction

Figure 16. A conceptual model of porous carbon. This is a three-dimensional representation of Figure 1a. Reprinted with permission from [46], R. C. Bansal et al., in *Active Carbon,*" Chapter 3. Dekker, New York, 1998. © 1988, Dekker.

between the gases (relative rates; $C + H_2O$: 3; $C + CO_2$: 1; $C + O_2 : 1 \times 10^5$) [48].

When carbon materials are exposed to activating gases, a two-stage pore formation process takes place reflecting the reactivity of the phases. In the first step the disorganized carbon is burnt out preferentially when the burn off does not exceed 10%. This results in the opening of the blocked pores. In the second stage the carbon of the aromatic ring system starts to burn, producing active sites and wider pores. In this way, a very complex and interpenetrating pore system is developed in the carbon materials. As the manner discussed previously, the pore size distribution is a function of the degree of activation. It is said that micropores are formed when burn-off is less than 50%, and macropores are introduced when burn-off is larger than 75%.

Physical activation can be accelerated and modified by the presence of metal elements. In the 1980s, Baker et al. conducted direct observations of catalytic gasification by using controlled atmosphere electron microscopy and found that transition metals made channels as they moved around the carbon matrices [49, 50]. This is the principle of conventional catalytic activation.

Oya et al. prepared the metal-containing ACFs, where the metal elements were introduced by mixing their complexes with phenol-formaldehyde resin [51, 52]. After spinning and stabilization, the raw fibers were subjected to carbonization and activation. Figure 17 shows the mesopore size distribution curves for the ACFs with manganese, zinc, copper, chromium, nickel, and iron. Introduction of Mn, Zn, and Cu gave no changes from the pristine fiber sample. Cr, Fe, and Ni showed various distribution curves (i.e., Cr enhanced the population of pores with 1–2 nm in radius, Fe formed mesopores spreading over the wide range of pore radius, and Ni gave a considerable development particularly in the 10–100 nm radius region).

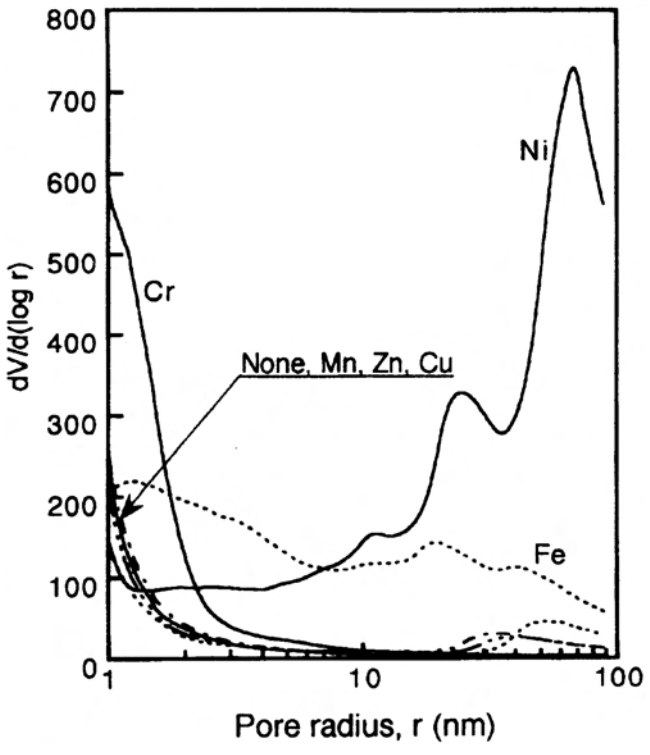


Figure 17. Mesopore size distribution curves of metal-containing activated carbon fibers. ACFs were prepared by spinning of metal compounds containing phenol-formaldehyde resin followed by carbonization and steam activation. Reprinted with permission from [52], J. Ozaki and A. Oya, *HYOMEN (Surface)* 35, 535 (1997). [in Japanese]. © 1997, Koshindo.

Tamai et al. reported that introduction of yttrium complex to a pitch finally gave mesoporous carbons [53, 54]. These carbons possess exclusively mesopores at a high rate such as 80% on surface area basis. They suggested the possibility of using these carbons for trapping large molecules such as vitamin B_{12} and humic acids in water. The formation of mesopores was attributed to the catalytic activity and to the movement of yttrium particles in the carbon matrix.

As seen the catalytic effects and the pore-size distribution are strongly dependent upon the kind of metal elements used as catalyst. Catalytic activation basically utilizes catalytic gasification, so that the differences in the effects should be attributed to the catalytic activity of each metal element. For further control of the pore size distribution, we have to understand the mechanism of catalytic gasification from the viewpoints of pore formation processes.

Metal elements can also be introduced by cation exchange. Nakagawa et al. studied the carbonization of styrene-based polymers with cation exchangeable sulfonic groups [55]. They introduced zinc, copper, calcium, and nickel into the polymer by cation exchange. When the exchanged cation was proton, the polymer shrank during carbonization at 1173 K and finally gave no pores in the carbon. In the cases of divalent or trivalent exchanged metals, carbonization resulted in the formation of micropores ranging 0.33–0.55 nm. The introduced cations have two roles in the pore formation. In the earlier stages of carbonization, the divalent or trivalent cations act as a bridging agent to resist shrinking against pyrolysis of the resin. Pores formed in these stages reinforced by metal-sulfide pillars as depicted in Figure 18. There seems to be an inconsistency between the pore size and the crystallite size of the metal sulfide crystallites.

Blending of iron compounds such as ferrocene with thermosetting resins like phenol-formaldehyde resin or furan resin resulted in the formation of mesoporous carbons. Ozaki et al. are interested in the microstructure and the electrochemical properties of the iron-containing carbons prepared from these mixture [56–59]. Figure 19 shows an electron micrograph characteristic of the iron-containing carbons obtained. Several tens of nanometers of iron particles were surrounded by partially ordered graphitic layers. Carbons containing such structures showed electrocatalytic activity comparable to a platinum electrode. Nitrogen adsorption study revealed that these iron-containing carbons included mesopores [59]. These mesopores may originate from the voids between the spheres shown in Figure 19.

3.1.3. Chemical Activation

Chemical activation includes degradation of raw materials with activating chemicals, such as zinc chloride, phosphoric acid, calcium chloride, etc. This method is usually applied to

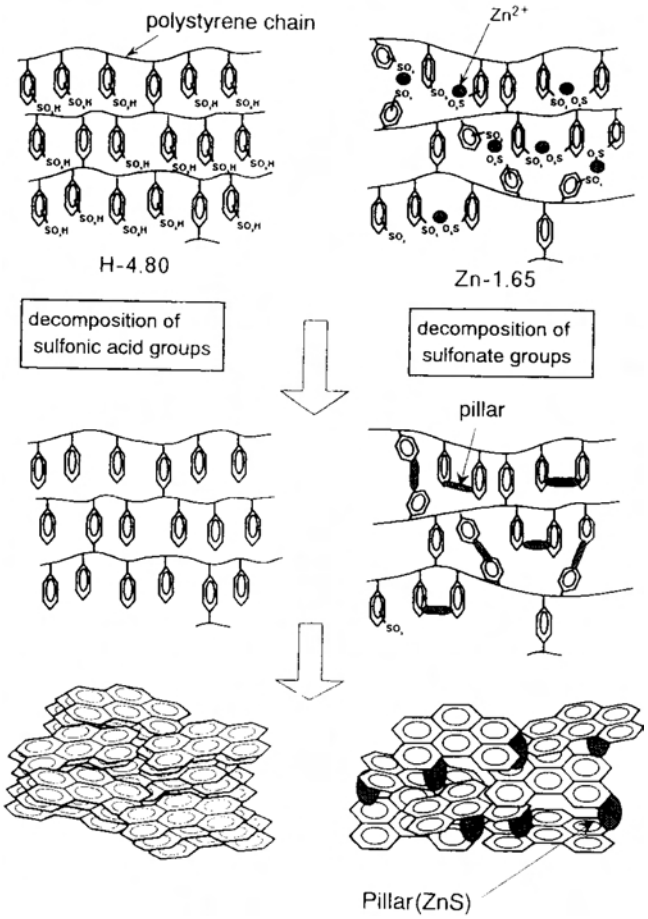


Figure 18. A mechanism for preparation of porous carbons from ion-exchange resins. The pillars are responsible for the formation of pores (right). Reprinted with permission from [55], H. Nakagawa et al., *Carbon* 37, 1455 (1999). © 1999, Elsevier.

wood origin materials, for example, sawdust [47]. Raw materials are impregnated with these chemicals and then heat-treated at elevated temperature ca. 673–1073 K. The main role of the activating reagent is to dehydrate and oxidize the chemical structure of the raw material. Such chemical alteration of the raw materials results in the suppression of tar formation and the promotion of carbonization. The temperatures required for chemical activation are normally lower than required for physical activation and hence we can obtain well-developed pores by using this method without worrying about pore shrinking at higher temperatures. In addition the pore size distribution strongly depends on the activator loading. Because of difficulties in scale-up and wastewater treatment, the chemical activation method is used only in restricted application fields.

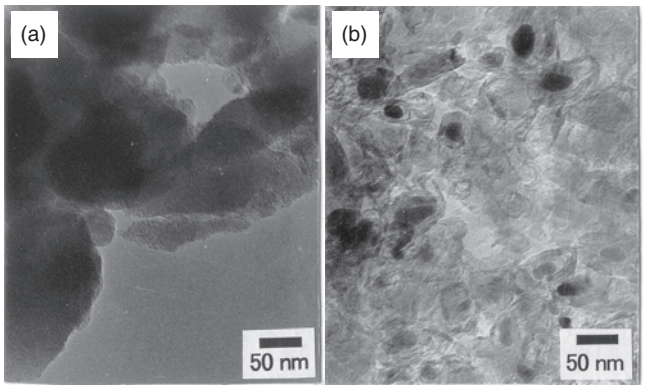


Figure 19. Transmission electron micrographs of the carbons (a) from pristine poly(furfuryl alcohol) and from a copolymer of vinyl ferrocene and poly(furfuryl alcohol).

Kadlec et al. reported that the shapes of the pore were different between the activated carbons obtained by physical and chemical activation processes [60]. The pores were usually bottle shaped in the case of chemical activation and cone shaped in physical activation. The bottle-shaped pores were attributed to the plastic state of the carbonizing materials, because of the low temperature activation (ca. 773 K). The gases released by thermal decomposition make hollows in the plastic material. When the gases escape from the hollow, small pathways would be created. That is why the chemical activation brings about a bottle-shaped pore. In the case of steam activation, there is a concentration gradient between the entrance of the pore and the central part of the pore. The gasification rate is higher at the pore entrance than at the pore center. Thus the conelike shape pore is likely created.

Activation by alkali metal salts should also be classified as a chemical activation. Heat treatment of raw materials mixed with excess amounts of alkali metal salts can provide activated carbon with specific surface area as high as 1000–4000 m^2/g. Weunerberg and O'Grady prepared activated carbons with high surface area from coal, coke, petroleum coke, and their mixtures by impregnating them with potassium hydroxide [61]. The mixture was first subjected to heat treatment between 573 and 723 K, and then to further heat treatment between 973 and 1123 K. The heat-treated samples were washed with water to remove the alkali metal, so that they gave activated carbons with surface area of 3000–4000 m^2/g for coke and of 1800–3000 m^2/g for coals. Schafer prepared activated carbons from a Victorian brown coal by adding potassium hydroxide [62]. The surface area of the activated carbon obtained was 1100–1500 m^2/g. Introduction of potassium hydroxide causes formaion of micropores, and the micropore volume is found to be a function of the potassium hydroxide loaded. For example, addition of 70% of KOH by weight to coal resulted in an increase in micropore volume by four times.

3.2. Novel Preparation Methods

Interest in mesoporous carbons has been growing recently instead of the traditional microporous carbons. This is probably for two reasons. One is increasing demand for the applications concerning liquid media, for example water treatment, EDLCs, and polymer electrolyte membrane fuel cells as described in Section 4. Another is interest in the preparation of ordered mesoporous materials stimulated by templated mesoporous silica.

Following are possible ways to prepare porous carbon materials stated by Ryoo et al. in their review [63]: (1) Deep activation, (2) utilization of metal catalysts, (3) coupling of physical activation and chemical activation,

(4) template methods, (5) carbonization of polymer gels, (6) polymer blend methods, and (7) multiwalled carbon nanotubes (MWNTs). The first three methods should be classified as conventional methods, as already mentioned. Of course physical activation still occupies an important position in practical meaning, so that much fundamental research on this subject is undertaken and published [64–67]. We will concentrate only on new preparation methods other than conventional physical or chemical activation methods. Taking into account the classification of Ryoo et al., the novel techniques were divided into five categories: (1) template methods, (2) polymer blend methods, (3) organic gel methods, (4) new methods for pore formation, and (5) utilizing MWNTs.

3.2.1. Template Methods

Template methods essentially consist of the following three steps: charging of organic materials into nanospaces of the template inorganic materials, carbonization of the organic–inorganic nanocomposites, and liberation of carbons by removing the template inorganic materials [68]. Knox et al. prepared mesoporous carbons by carbonization of phenolic resin impregnated silica gel or porous glasses followed by removing the inorganic templates [69–71]. This carbon is commercially available for column packing materials for liquid chromatography. Kamegawa and Yoshida also carbonized organic–silica composite and obtained carbon materials that can be swollen by contact with organic vapors [72, 73]. In this case, they made very thin organic layers that were fixed by estrification to the surface of template silica.

In the previous examples, the pores in the porous materials were utilized as templates, in other words, carbon precursors penetrated into the template's pores. There are examples including embedded templates. Hyeon et al. used a colloidal silica (ca. 12 nm in diameter) as a template to be embedded in the resorcinol resin [74]. They obtained carbons of which surface area, total pore volume, and mesopore volume were 950 m^2/g, 5.5 cm^3/g, and 4.8 cm^3/g, respectively, when the resorcinol/silica ratio was 7.5. Li and Jaroniec developed an imprinting method in which template silica particles were imprinted into mesophase pitch spheres [75]. In this case, carbonization and template removal were also included. This method gave a possibility of controlling the spatial distribution of the pores such as uniform distribution or biased distribution to surfaces depending on the extent of the imprinting depth into the mesophase pitch, as depicted in Figure 20. Kawashima et al. prepared mesoporous carbons from polymerized furfuryl alcohol in the presence of silica sol made by hydrolysis of tetraethoxysilane [76].

There are many attempts to use ordered inorganic materials as templates. Clays and zeolites form a class of porous materials with ordered pore structures defined by crystalline ordering. Kyotani et al. synthesized very thin carbon film by carbonization of the *in-situ* polymerized polyacrylonitrile or poly(furfuryl alcohol) in the interlayer space of a montmorillonite [77–79]. The obtained carbon was easily graphitized even though the carbon precursors were polymers that usually gave nongraphitizing carbons. Actually, this is the starting point of this type of study. Kyotani et al. tried to prepare carbon material with controlled pore distribution by using these microporous crystalline templates [80]. They could obtain high surface area carbon (2200 m^2/g) by using Y-zeolite as a template [81]. They also made another effort to fill the zeolite's pores with carbon as much as possible by utilizing both *in-situ* polymerization of furfuryl alcohol and chemical vapor deposition of propylene into pore spaces. The obtained microporous carbon had a periodical ordering of 1.4 nm, which is the same as the spacing of the 111 plane of Y-zeolite as shown in Figure 21.

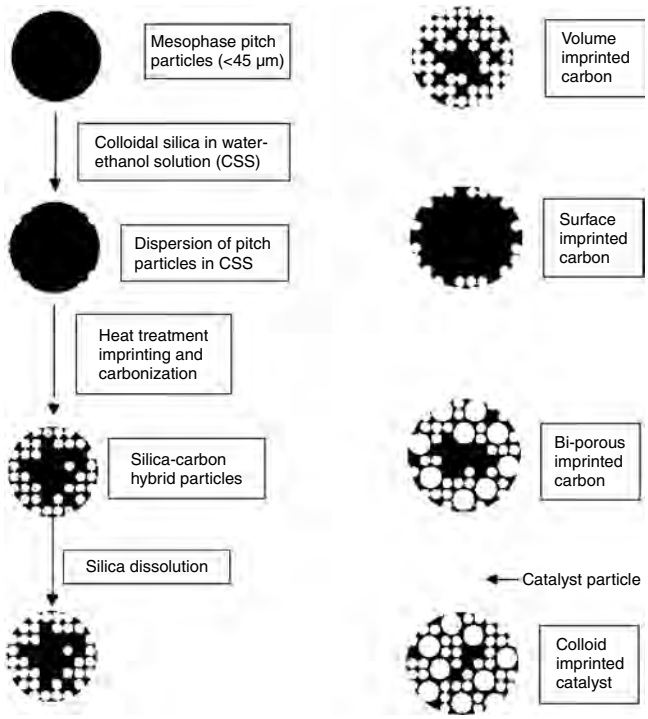


Figure 20. Concept of colloid imprinting method for preparing spatially regulated porous carbons. Reprinted with permission from [75], Z. Li and M. Jaroniec, *J. Am. Chem. Soc.* 123, 9208 (2001). © 2001, American Chemical Society.

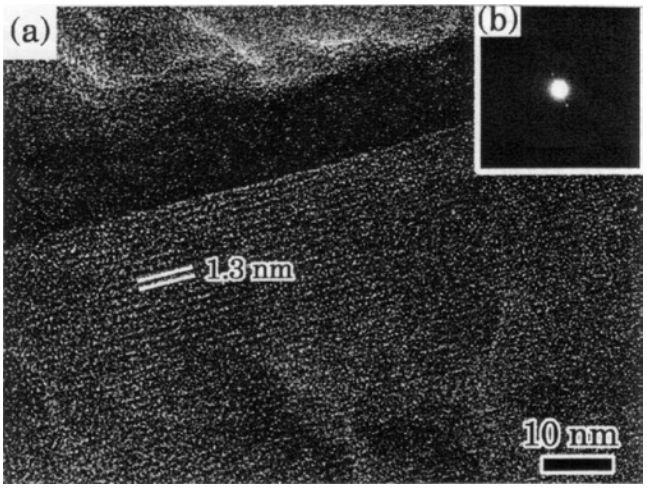


Figure 21. (a) High resolution transmission electron micrographs of a carbon prepared by using Y-zeolite as a template and (b) its electron diffraction pattern. Reprinted with permission from [81], Z. Ma et al., *Chem. Commun.* 2365 (2000). © 2000, Royal Society of Chemistry.

Apart from the utilization of microporous inorganic material as template, two Korean research groups reported that they successfully prepared regularly ordered mesoporous carbons by using mesoporous silica as templates. Lee et al. used an aluminum substituted mesoporous silica, MCM-48, as a template [82]. Phenol and formaldehyde were polymerized in the template and subjected to carbonization and removal of template. The obtained carbon showed an X-ray diffraction pattern resembling the template. Electron microscopic analysis revealed that there were small pores of 2 nm in diameter separated by a carbon wall with a thickness of 2 nm. Ryoo et al. also reported the formation of highly ordered porous carbon by using the same mesoporous silica, MCM-48 [83]. However, they employed different carbon precursor materials, sucrose and sulfuric acid. Figure 22 shows the X-ray diffraction of the template (a), template with carbon (b), and carbon itself (c). Highly ordered pores can be observed in an electron micrograph presented in Figure 23. Recently Joo et al. compared the carbon precursor for the mesoporous silica templated carbon [84]. When furfuryl alcohol was used as a precursor, rodlike carbon was obtained, while in the case of sucrose with sulfuric acid, like Ryoo et al., pipelike carbon was obtained. They loaded platinum on the pipelike carbon and found that the activity for oxygen reduction in fuel cells was extremely high (100 A/g-Pt).

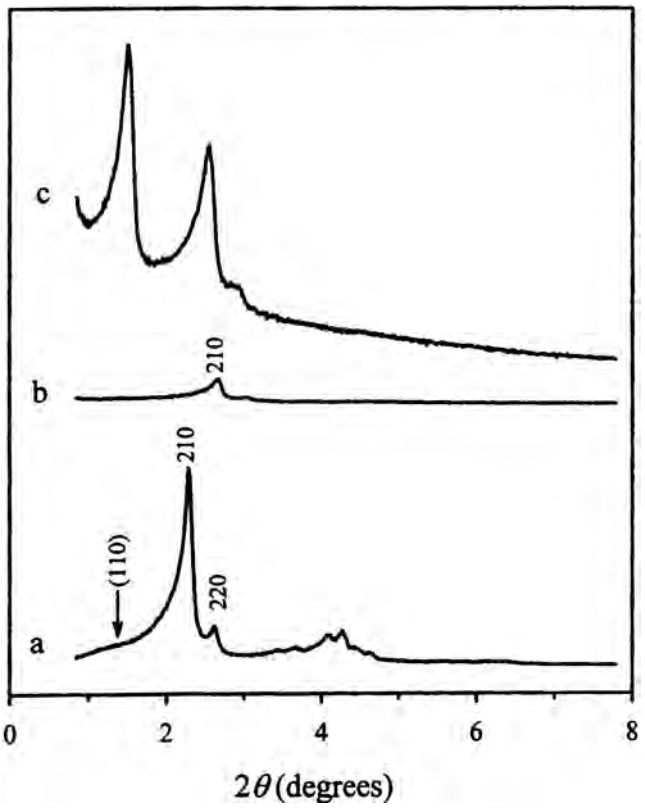


Figure 22. X-ray diffractograms of (a) MCM-48 (template), (b) MCM-48 after completing carbonization within pores, and (c) carbon obtained by removing silica wall after carbonization. Reprinted with permission from [83], R. Ryoo et al., *J. Phys. Chem. B* 103, 7743 (1999). © 1999, American Chemical Society.

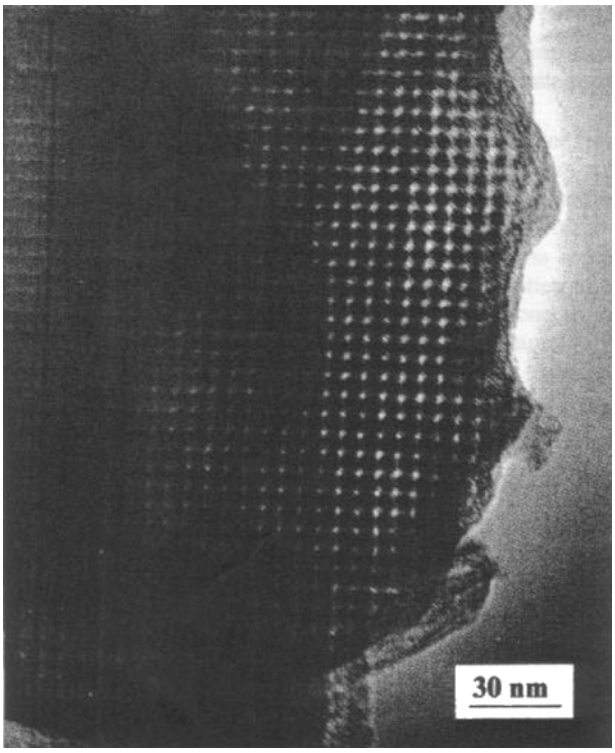


Figure 23. A high resolution transmission electron micrograph of a carbon prepared using MCM-41 as a template. Reprinted with permission from [83], R. Ryoo et al., *J. Phys. Chem. B* 103, 7743 (1999). © 1999, American Chemical Society.

Alkaline metal amalgam or alkali metal can defluorinate poly(tetrafluoroethylene) (PTFE) to produce carbynelike structure (*sp*-carbon allotrope) and finely dispersed metal fluoride crystallites in the polymer matrix. Washing out the metal fluoride from the matrix promoted a transformation from *sp*-hybridization to sp^2-hybridization by a cross-linking reaction between the adjacent carbyne chains [85, 86]. Shiraishi et al. found that heat treatments after the reaction between PTFE and alkali metals can alter the pore size distribution of the resultant carbons, (i.e., higher heat treatment temperature above 623 K caused a formation of mesopores) [87, 88]. This was attributed to the crystalline growth of the alkali halides that are anticipated to be templates for pore formation.

Fenelonov et al. reported a unique example of the template method [89]. They used carbon black as a template and deposited pyrocarbon from C_1–C_4 gases at 1123–1223 K. Templates were removed by steam gasification. In this process, at first the most defected parts of pyrocarbon were gasified. When the formed pore reached the carbon black, a preferential gasification started as shown in Figure 24. A distinct feature of this carbon was high resistance to attrition and mechanical crushing. They anticipate that this carbon can be applicable to catalyst supports.

Although the obtained pore was classified as macropores, Zakhilov et al. prepared spatially ordered porous carbon material by using opal as template [90]. Figure 25 shows the micrographs of the obtained carbon material.

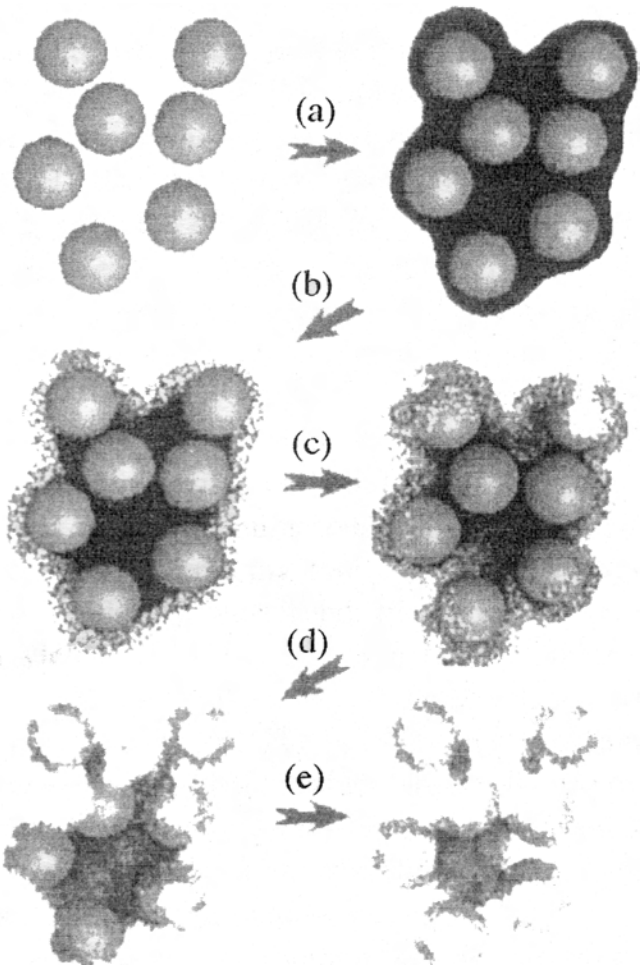


Figure 24. Schematic diagram for preparing porous carbons using carbon particles as template. Reprinted with permission from [89], V. B. Fenelonov et al., *Catal. Today* 42, 341 (1998). © 1998, Elesevier.

3.2.2. Polymer Blend Method

The polymer blend or polymer alloy method has been developed as a way to improve properties of polymers by blending other polymers. In this sense, the polymer blend is a kind of composite material. Polymer blends are divided into two categories depending on mixing levels. One is macroscopic blending and another is microscopic blending. Phase separated polymers are usually included in the first category and blends with compatibilizers, block co-polymers, and grafted co-polymers are included in the second category.

Pyrolysis behaviors of polymers are decided by the competition between scission of the polymer main chain and formation of cross-linking. Polymers with a tendency to main chain scission are unstable against heat treatment and therefore such polymers decompose without leaving carbon residues. Here these kinds of polymers will be referred to as pyrolyzing polymers. On the other hand, polymers with a tendency to form cross-linking can stand with the heat treatment and consequently leave carbon materials. Usually the latter category polymers are used as carbon precursors, and tentatively we call such polymers carbonizing polymers in this chapter.

Let us consider polymer blends between the pyrolyzing and carbonizing polymers. When we heat such a polymer blend for example up to 1273 K, theoretically we will have a porous carbon without further activation process, by leaving pores where the pyrolyzing polymer existed in the carbon matrix from the carbonizing polymers. If we control the distribution of the pyrolyzing polymer in the carbonizing polymer matrix, we will obtain carbon materials with controlled pores in them. This is the principle of the polymer blend method for preparing porous carbon materials as illustrated in Figure 26 [91].

Hatori et al. prepared a carbon from a polymer blend between polyimide and polyethylene glycol [92, 93]. In this

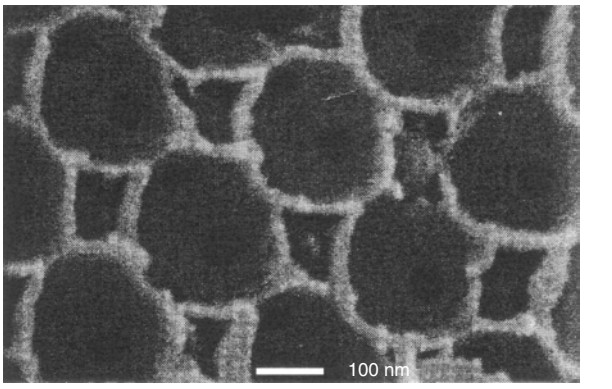


Figure 25. Regularly ordered macroporous carbons by using opal as a template (SEM image). Reprinted with permission from [90], A. A. Zakhilov et al., *Science* 282, 894 (1998). © 1998, American Association for the Advancement of Science.

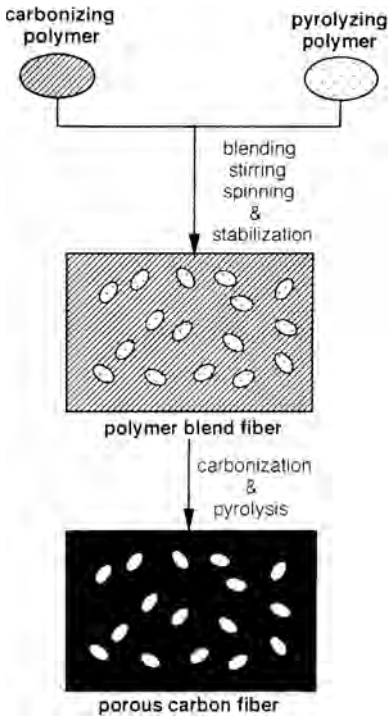


Figure 26. A concept of polymer-blend method for preparing porous carbons without activation. Reprinted with permission from [91], J. Ozaki et al., *Carbon* 35, 1031 (1997). © 1997, Elsevier Science.

case, polyimide was the carbonizing polymer and polyethylene glycol was the pyrolyzing polymer. They successfully obtained mesoporous carbon of ca. 5 nm diameter. Takeichi et al. used co-polymers between amide and urethane [94]. In this case the former was a carbonizing and the latter was a pyrolyzing polymer, respectively. The carbon material obtained from this polymer was indeed porous; however, the pore size was as large as 10 μm. These could be classified as macropores.

Ozaki et al. tried to use the polymer blend method for preparing ACFs [91]. In this case, the materials for the blends were limited from the standpoints of spinnability, compatibility, and difference in the decomposition temperatures between the selected polymers. They chose novolac-type phenol-formaldehyde resin as the carbonizing polymer and poly(vinyl butyral) as the pyrolyzing polymer. Figure 27 shows the weight decrease during heat treatment. The pyrolyzing polymer [poly(vinyl butyral)] shows a steep decrease in weight between 573 and 673 K. The carbonizing polymer (phenol-formaldehyde resin) showed a stepwise weight decrease. The first weight decrease is observed at around 523 K followed by a greater decrease between 673 and 823 K than the former. The final weight loss of this sample was 40%. For a 1:1 polymer blend, the final weight loss was 73%, which is very close to the simple average of the weight losses of both component polymers. Pore size distributions for micropores and mesopores were obtained by N_2 adsorption and are presented in Figure 28. As can be seen, the carbon fiber made from the polymer had a suppressed micropore volume and an enhanced mesopore volume. However, the obtained increase in the pore volume was less than they expected. Ozaki et al. analyzed the details of the polymer blend by using Fourier transform infrared and TG-MS techniques [95]. They found that when the

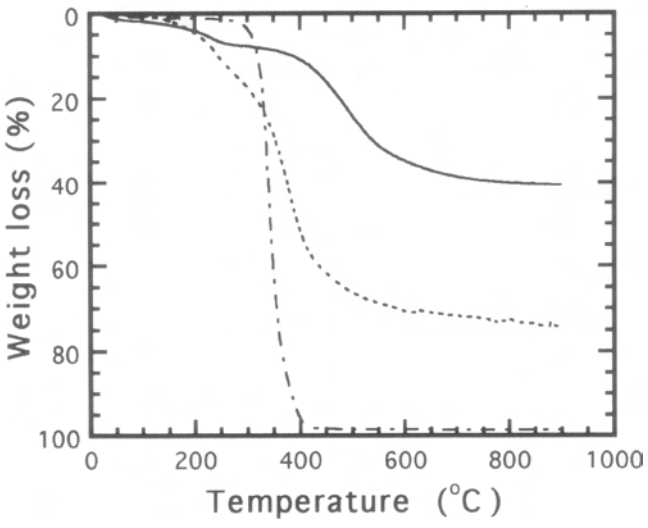


Figure 27. Weight loss curves for a polymer blend and its component polymers. Broken line: polymer blend; dash–dot line: pyrolyzing polymer; solid line: carbonizing polymer. The polymers used here were a 1:1 blend of phenol–formaldehyde resin (carbonizing polymer) and poly(vinyl butyral) (pyrolyzing polymer). Reprinted with permission from [91], J. Ozaki et al., *Carbon* 35, 1031 (1997). © 1997, Elsevier Science.

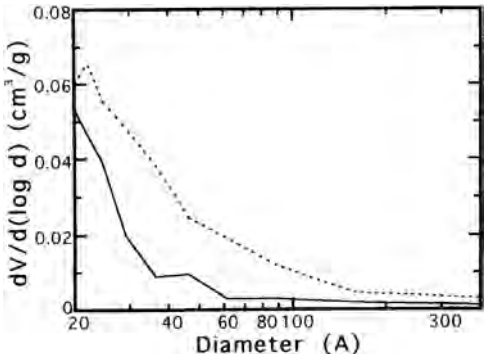


Figure 28. Mesoporous size distribution of the carbons prepared from a phenol–formaldehyde resin (solid line) and from a poly(vinyl butyral) (broken line). Reprinted with permission from [91], J. Ozaki et al., *Carbon* 35, 1031 (1997). © 1997, Elsevier Science.

two components were prepared, they chemically interacted, which inhibited the independent pyrolysis of each polymer resulting in insufficient pore development.

Yang et al. prepared size-controlled mesoporous carbon spheres from polymer blends of novolac-type phenolic resin and two pyrolyzing polymers, poly(ethylene glycol) (PEG) and poly(vinyl butyral) (PVB) [96]. They put only a small amount of pyrolyzing polymers (7.5 units of PVB and 15 units of PEG were blended to 100 units of phenol-formaldehyde resin on weight basis), and finally they subjected the carbonized materials to steam activation. The obtained activated carbons possess mesopores ranging 3–5 nm. Addition of ferrocene to the polymer blend could give additional larger pores (10–90 nm) to the polymer blend derived activated carbons.

If a pyrolyzing polymer includes metal precursor and is dispersed in carbonizing polymer matrix, we can expect a porous carbon material on which pore surfaces are deposited by metal particles. Ozaki et al. prepared platinum loaded carbon fibers with the same polymer combination as described previously [97, 98]. By comparing the platinum loaded carbon fiber with platinum loaded fibers prepared by a simple mixing method, the former gave a higher hydrogen adsorption ratio and showed platinum particles deposited exclusively on the pore wall by electron microscopic observation.

During carbonization unstable parts in carbon precursor materials are eliminated as volatile matters. Devolatilization can introduce voids or pores in the resultant carbons. Such a preparation technique can be included in the polymer blend method, as the raw polymers are designed to have carbonizing parts and pyrolyzing parts. Takakura et al. synthesized a series of triple bonds containing polymers, poly(phenylene butadiyne) derivatives as shown in Figure 29, and studied their pore structures [99]. The micropore volume was around 0.3 ml/g independent of the polymer; however, the mesopore volume varied from 0.01 to 0.47 ml/g. The highest mesopore volume was achieved on P3 structure in Figure 29. Furthermore, the mesopore size distribution was tremendously influenced by the kinds of side chain groups. They considered that mesopores were introduced by the elimination of side chain groups. There is another example of preparing porous carbons by modifying the original polymer structure. Lenghaus et al. prepared phenolic resin from

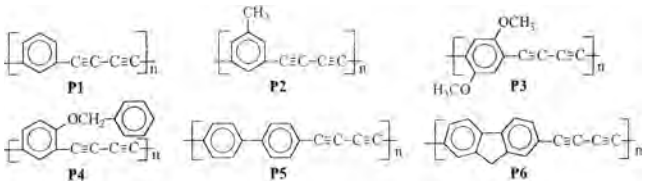


Figure 29. Structurally modified poly(phenylene butadiyne) polymers for carbonization. Reprinted with permission from [99], T. Takakura et al., in "Proceedings of the 28th Annual Meeting of the Carbon Society of Japan," Kiryu 2001, p. 208 [in Japanese] © 2001, Carbon Society of Japan.

alkyl-substituted phenols and studied the influence of these groups on the formation of porous structures [100]. By comparing the CO_2 and N_2 surface areas, para-alkyl substituted phenols gave wider micropores. Thus changing the substitution group will be a promising way of controlling the pore size, but we need to know the carbonization mechanisms of the parent polymers for designing the raw materials.

3.2.3. Sol–Gel Method

Preparation of mesoporous carbons by carbonizing organic gels was initiated by Pekala [101]. In his study, resorcinol was condensed with formaldehyde in alkaline solution. A schematic diagram of the resorcinol–formaldehyde (RF) gelation mechanism is presented in Figure 30. Resorcinol was first substituted with hydroxymethyl group by a catalytic action of alkaline reagent. The products condensed into surface functionalized polymer "clusters" that cross-link to form a gel. Pekala and co-workers carbonized the organic gel and found that the resulting carbon aerogel has high porosity (>80%) and a high surface area (400–900 m^2/g) [102]. The mesoporosity of the carbon aerogels can be ascribed to a network structure that was inherited from the original organic aerogel.

Tamon et al. conducted exhaustive studies on the relation between the porous nature of the carbon aerogels and their sol–gel conditions and showed that the pore size and the pore volume can be controlled by changing the mixing ratios of resorcinol, formaldehyde, water, and catalyst [103]. They tried to use freeze drying instead of supercritical drying for drying wet gels, because the former is more economical than the latter [104]. Hanzawa et al. conducted high temperature treatment of the carbon aerogel [105]. They found that micropores disappeared by 2273 K, but half of the mesopores could survive even after a heat treatment at 3073 K.

Resorcinol is not the only reagent for the preparation of organic gel. Gel formation was also observed for phenolic–furfural [106], melamine–formaldehyde [107], polyurethane [108], and mixed cresol–formaldehyde [109] systems. Finding cheaper raw materials and developing a cost-effective and time-saving drying method should be required for the practical uses of sol–gel carbons.

The application fields of carbon gels will be expanded, because the gels can be formed into desirable forms such as monoliths, wafers, films, spheres, and irregular powders [110]. In particular, monolith or film type carbon gels are suitable for electrochemical usage such as electrodes for EDLC [111, 112], for capacitive deionization [113], porous electrodes for fuel cells [114], and analytical electrodes [115].

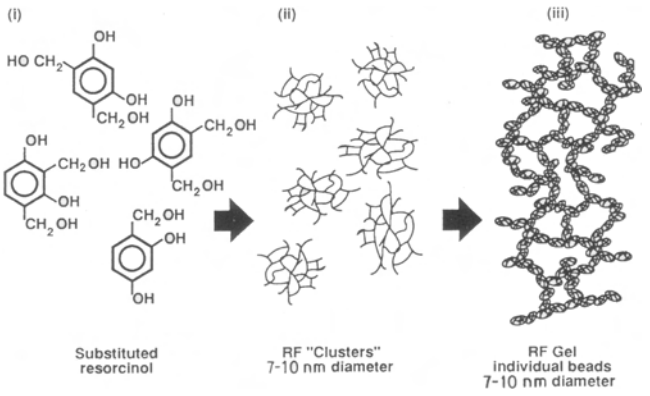


Figure 30. Formation of a silicalike organic gels from resorcinol. Reprinted with permission from [101], R. W. Pekala, *J. Mater. Sci.* 24, 3221 (1984). © 1984, Kluwer Academic Publishers.

3.2.4. Unique Pore-Forming Methods

In the traditional way to prepare activated carbons, pores are introduced by exposing carbons to oxidizing gases such as steam or carbon dioxide at elevated temperatures. Although the studies are not so many, there are unique activation methods such as electrochemical, supercritical fluid (SCF), and microwave assisted oxidation.

Electrochemical oxidation of carbon fibers has been employed to improve the shear properties of the fibers. Pittman et al. studied electrochemical oxidation of carbon fibers using the continuous operating system [116–118], as represented in Figure 31. The electroxidation apparatus permitted running carbon fiber tows through the U-tube using a motor and a set of gears. Carbon fiber from the incoming spool contacts the first rotating metal cylinder before going

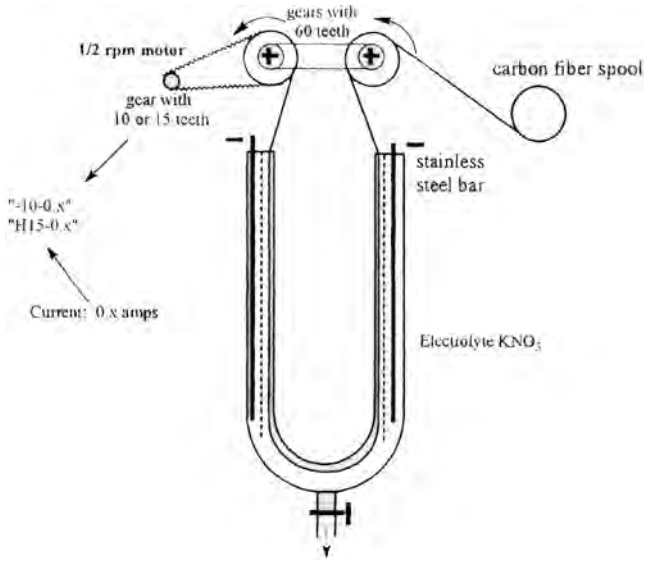


Figure 31. An apparatus for pore introduction to a carbon fiber by a continuous electro-oxidation of carbon fibers. Reprinted with permission from [116], C. U. Pittman, Jr., et al., *Carbon* 37, 85 (1999). © 1999, Elsevier Science.

through the entire U-tube and contacting the second rotating metal cylinder, where it is rewound. These two metal cylinders serve as the positive electrode. The two stainless steel bars in the U-tube act as the negative electrode. They used 1 wt% KNO_3 solution as electrolyte. By using adsorption, NaOH titration, and X-ray spectroscopy (XPS) techniques, electro-oxidation was an activation method that could introduce highly interconnected ultramicropores extending far below the depth probed by XPS.

SCFs have gaslike and liquidlike properties, such as high diffusion coefficients and low viscosities similar to gases, and excellent dissolving abilities. SCFs have now been used in extraction and separation fields. Recently, their unique chemical properties attracted people's attention for recycling plastics [119, 120] and decomposition of environmentally harmful substances. Applications of SCFs to materials processing are now studied, for example particle formation by rapid expansion of supercritical solution [121] and antisolvent deposition [122]. Li et al. conducted a study examining the feasibility of using supercritical water or carbon dioxide as the activating agent for carbon materials [123]. They employed vapor grown carbon fiber as the carbon material to be activated. After the treatment at 653 K and 250–320 kgf/m^2, surface areas of the carbon fibers were slightly increased from 9.29 to 15.85 m^2/g for water and to 13.64 m^2/g for carbon dioxide. Taking into account the critical points of water (T_c: 647 K, P_c: 22 MPa) and carbon dioxide (T_c: 304 K, P_c: 7.3 MPa), the treatments were done in the supercritical conditions of each substance. An enhanced effect was obtained for both media, if the carbon fibers were pretreated with nitric acid (i.e., they observed 35.27 m^2/g for water and 33.66 m^2/g for carbon dioxide). Adsorption abilities of the ACFs by means of supercritical treatments were studied for benzene and phenol, but a commercially available ACF showed the highest ability.

As microwave heating can give internal and volumetric heating, microwave-induced reaction is expected to proceed more quickly and effectively at a lower bulk temperature. Chars are reported to be a good receptor of microwave energy and may reach the minimum reaction conditions. Guo and Lau conducted a study on microwave induced CO_2 activation of an oil-palm-stone char [124]. A microwave power of 450 W was the minimum power required to cause any change from the viewpoint of volatile content and adsorptive properties such as BET surface area and pore volumes. Addition of CuO to the char induced a further activation resulting in the formation of a mesopore region, because the substance acted as a microwave receptor to increase the effective activating temperature.

Usually porous carbon material has inferior electrical conductivity to nonporous carbon materials produced at the same temperature. Lu and Chung developed a high surface area with high conductivities [125]. The idea of the preparation is to include a graphitization process. They examined the following methods to produce such carbons from PAN-based carbon fibers (specific surface area: 7 m^2/g), pitch-based carbon fibers (5 m^2/g), and vapor grown carbon fibers (54 m^2/g): (1) activating graphitized carbon after ozone treatment and (2) activating carbon that has not been graphitized, followed by graphitization and second activation. Both methods could increase porosity for all of the samples used here. In the case of method (1), the surface areas of the samples were ca. 400 m^2/g with resistivities of (2–4) $\times 10^{-3}$ Ω cm. On the other hand, method (2) produced porous carbon materials with 200–300 m^2/g with resistivities of (2–4) $\times 10^{-3}$ Ω cm. From the viewpoints of surface area development and economic merit, they favored method (1) for producing porous carbon materials with high conductivity.

3.2.5. Nanocarbons

This nomenclature has been widely used since an international symposium on nanocarbons was held in Nagano, Japan in November 2002. Inagaki and Radovic interpret it in terms of the structural and textural features [126]. They summarize the controlling factors for the production of various classic and new carbon materials containing the nanocarbons, as indicated in Figure 32. In regard to the pore structure, they explain that what are traditionally called microporous carbons in which the dominant pores are the so-called micropores, <2 nm, are increasingly being referred

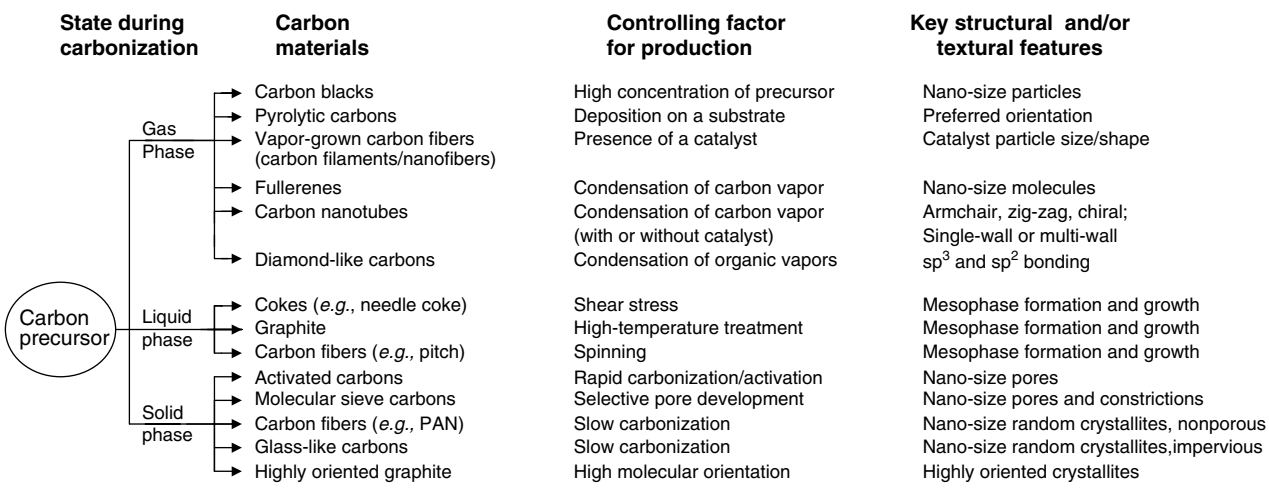


Figure 32. Controlling factors for the production of various carbon materials and the key points for their structural and/or textural features. Reprinted with permission from [126], M. Inagaki and L. R. Radovic, *Carbon* 40, 2263 (2002). © 2002, Elsevier Science.

to as "nanoporous" carbons, because the majority of the pores are of nanometer size. Also the category of nanocarbons is proposed on the basis of nanometer-scale structure and size of newly developed carbons, as illustrated in Figure 33. In this section we focus only on the carbon nanotubes in the nanocarbons.

Carbon nanotubes and their related materials are forming a new category of nanoporous materials. In 1997, Dillon et al. first clamed that SWNTs have a high reversible hydrogen storage capacity [127]. It was the start of the following a large number of studies on hydrogen storage. Here we will not review these studies, but we explain the overview of these materials from the viewpoint of pore structure. SWNT is a carbonaceous tube with a diameter of 1 nm and usually makes hexagonally packed bundles (Fig. 5). Figure 34 shows a configuration of H_2 molecules on carbon nanotube arrays by a Monte Carlo calculation [128]. It illustrates that H_2 molecules adsorb in the interstitial cavity and in inner-tube cavities. The attraction potential inside the tube is calculated to be larger because the curvature of the tube increases the number of nearest neighbor carbon atoms.

Eswaramoorthy et al. were interested in the adsorptive nature of carbon nanotubes without worrying about the surrounding graphitic layers [129]. So they conducted a study by selecting a SWNT rather than using a multiwalled carbon nanotube that was known to possess mesopores [130]. They prepared HNO_3-treated SWNTs as well and compared the adsorption nature of N_2 and benzene. The fact that the HNO_3-treated SWNT showed a larger amount of benzene and the calculation of the cross-sectional area of benzene lead them to conclude that the inner-tube space was also used for adsorption.

Barisci et al. studied the electrochemical nature of SWNTs from the standpoint of electrochemical capacitance [131].

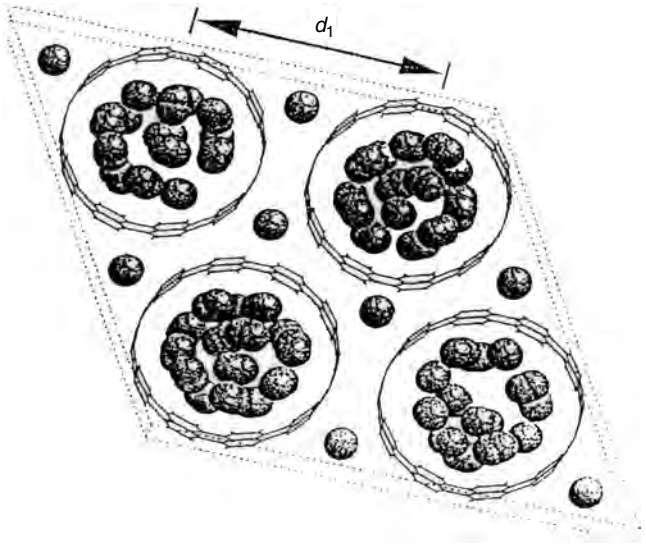


Figure 34. Configuration of hydrogen molecules in a bundle of single-walled carbon nanotubes. Reprinted with permission from [150], H. M. Cheng et al., *Carbon* 39, 1447 (2001). © 2001, Elsevier.

The capacitance of the nanotube paper (NTP) made from SWNTs was independent of the sizes of the cation or anion and of the voltage sweep rate for the measurements. These results indicated that the NTPs investigated are characterized by an interbundle open structure with relatively large pores that can be readily accessed by ions with a wide range of sizes and charges. Possibly, this is due to the pores being interconnected spaces in the entangled nanotube network rather than cavities and micropores as in other carbon materials.

The adsorption nature of SWNTs is not so simple as is judged from its appearance. It is important to modify the

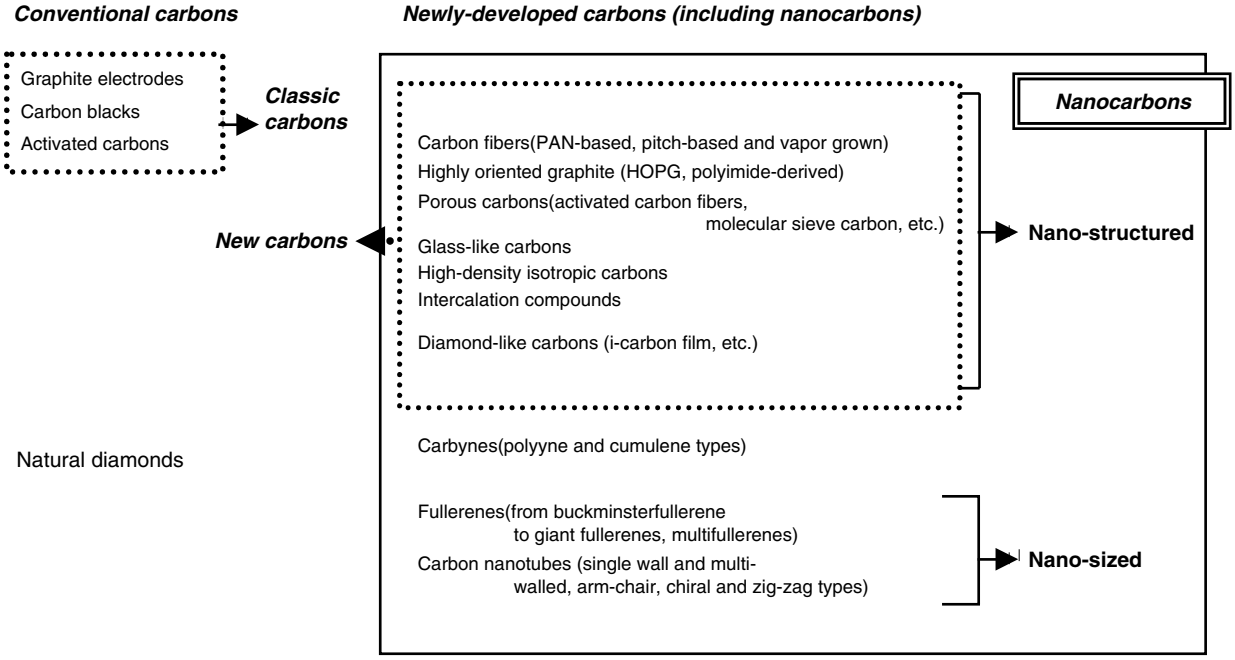


Figure 33. Proposed scheme for classification of carbon materials, including nanocarbons. Reprinted with permission from [126], M. Inagaki and L. R. Radovic, *Carbon* 40, 2263 (2002). © 2002, Elsevier Science.

higher order arrangement as well as the bundle structure and the inner-tube spaces depending on the applications of the SWNT. The situation of multiwalled carbon nanotubes seems more complex, because this substance is known to include mesopores.

4. APPLICATION OF NANOPOROUS CARBONS

4.1. Traditional Applications

Discoloration of sugar, air purification, removal of odor or toxic gases, water purification, gas separation, solvent recovery, etc. have been the traditional application areas of porous carbons. Most of these applications are based on the superior adsorption properties. Of course, the pore sizes of the activated carbons are different among the various applications by controlled production of activated carbons. Because the readers may find a suitable book for activated carbons that describes these historical usages of the activated carbons [132], we will review contemporary applications of porous carbons.

4.2. Nontraditional Applications

In recent days, we are facing many problems making our life dangerous and uncomfortable, such as pollution induced by highly toxic chemicals like dioxins or environmental endocrine disruptors (environmental hormones), air pollution caused by emitted gases from vehicles and industries, and shortage of energy and foods particularly in the developing countries. Political efforts are the main means to solve these problems; however, it is also important to provide techniques to back up the efforts. Porous carbon is now earning its good reputation as an agent to solve the problems due to developments in fine techniques and sciences to prepare the desired properties, as we have already learned in the previous section.

In this section, we will give a brief review of porous carbon application in the fields of gas storage, catalysts, removing harmful gases, electronics, and medicine.

4.2.1. Gas Storage

The target gases for storing are mainly methane and hydrogen.

Natural gas gathers attention as a clean energy source and is expected to be an automobile fuel [133]. Compression and liquefaction are the currently available techniques to store it and are called CNG (compressed natural gas) and liquefied natural gas (LNG), respectively. Gas is charged in a container at 25 MPa in CNG, so that the pressure resistance is required for the container and the valves, which lead to restricted application areas. In the LNG application, natural gas is liquefied at 111 K, which requires incidental facilities for lower temperatures. This technique is not also suitable for vehicle uses. ANG stands for adsorbed natural gas. In the ANG technique methane is usually adsorbed on a suitable adsorbent at a relatively lower pressure (3.5 MPa) at ambient temperature. The key for the practical uses of this technology is to find effective and cheap adsorbents. Activated carbons, porous polymers [134–137], silica gels, and zeolites [138–141] have been examined as candidates for the adsorbents.

As the critical point of methane is $T_c = 190.55$ K, $P_c = 4.595$ MPa, the substance is in the subcritical or in the supercritical states. Supercritical gases tend to selectively adsorb in the pores of which pore width is about 4–5 times the molecule size [133]. It is also well known that the equilibrium amount decreases with the increase in pore width. So the adsorbent should not have pores larger than 2 nm. Further the materials are required to be microporous with large density, because the adsorption ability is evaluated by the amount of adsorption per unit volume of the adsorbent rather than by the amount per weight of the material. The researchers are targeting the value of 150 V_m/V_a at 298 K and 3.5 MPa, where V_m and V_a stand for the volume of adsorbed methane and the volume of the adsorbent, respectively. The Gas Research Institute presented more severe conditions for practical automobile applications, 200 V_m/V_s (STP) and less than \$4.41 per a kilogram [142]. Preparing microporous carbon materials with high packing density at cheaper prices can satisfy these requirements. From a theoretical calculation, the adequate pore sizes are 1.12 [143] or 1.14 nm [144]. If the density of an adsorbent is 0.67 g/ml, which is a typical value for a monolithic carbon, the maximum methane uptake is estimated to be 220 V_m/V_a at 298 K.

The carbon materials that selectively possess micropores are ACFs. Alcaniz et al. examined ACFs with different pore characteristics prepared by different activating agents and conditions for methane storing ability [145]. They found that the methane uptake has a good relation between the micropore volumes and obtained the highest value such as 163 V_m/V_a for adsorption and 143 V_m/V_a for desorption of methane.

Coal, biomass, and used tires are cheaper raw materials for use in porous carbons for ANG. Sun et al. conducted a study with coal and found that KOH activated carbon has a larger methane uptake than the steam activated one on the weight basis, but the uptake by the steam activated one was larger than the KOH activated one, because of the smaller density of the KOH activated carbon [146]. Catula et al. found that the activated carbon made from peach kernel with zinc chloride activation showed a high adsorption ability for hydrocarbons [147]. MacDonald and Quinn extended the study to examine the methane storing ability by using phosphoric acid as an activator [148]. The methane storing ability was less than the master curve for other carbons. This was probably caused by the presence of abundant surface functional groups. Used tires seem to be a promising candidate for the adsorbents, but Muller et al. could obtain only a 47 V_m/V_a storing index [149]. As the utilization of waste materials is a very important subject of contemporary engineering, efforts like these must proceed.

Hydrogen must be the ultimate fuel for automobiles because it emits only water after combustion of the fuel. Whether the power generator is direct combustion or fuel cells, the common elemental technique is storing hydrogen. The US Department of Energy Hydrogen Plan has provided a commercially significant benchmark for the amount of reversible hydrogen adsorption [150]. The benchmark requires a system-weight efficiency (the ratio of stored hydrogen weight to system weight) of 6.5 wt% hydrogen and

a volumetric density of 62 kg-H_2/m^3. Three types of storing technology are considered: (1) cryogenic liquid hydrogen, (2) compressed gas storage, and (3) metal hydride storage technology. However, these approaches are not sufficient to satisfy the previous requirements [150]. Since Dillon et al. found the reversible hydrogen storage capacity for SWNT [127], many research groups started to conduct hydrogen storage experiments and have made some noticeable progress. The carbons examined for this approach were essentially so-called nanocarbons, like SWNTs, MWNTs, carbon nanofibers (CNFs), and nonstructured graphite. The amount of hydrogen adsorbed and the measurement conditions are listed in Table 3 [127, 151–165].

The results reported so far are varied and controversial due to poor or no reproducibility of the data. Recently David et al. measured the adsorption capacity using a variety of materials: carbon nanotubes, carbon aerogels, activated carbons, and so on. In the case of the nanotube samples all of the SWNTs and MWNTs tested showed poor adsorption values (less than 0.5 wt%) at 300 K and 10 MPa [166]. Also Takagi et al. reported the adsorption ability of hydrogen for ACF, a SWNT, and a zeolite [167]. In the study they improve the reliability of the data by confirming the reversibility between the adsorption and desorption at each equilibrium pressure. The values for the ACFs which showed a high adsorption amount in the samples, are around 0.2 wt% at 303 K and 3 MPa. At the moment, therefore, the capacitance level of hydrogen adsorption seems not to be so high, compared with the data listed in Table 3. In order to obtain accurate and reliable amounts of the adsorption, special attention should be paid to the preparation and purification of samples as well as the design of high pressure apparatus without leakage of hydrogen. Moreover, the adsorption mechanism of hydrogen in the nanoporous sites has to be elucidated for the improvement of the adsorption capacity.

Table 3. Summary of reported hydrogen storage capacities in nanocarbons.

Material	Amount of hydrogen storage (wt%)	Temperature (K)	Pressure (MPa)	Ref.
SWNT	5–10	133∼	0.04	[127]
SWNT	2–4	133∼	0.04	[151]
SWNT	4.2	R.T.	10–12	[152]
SWNT	0.9	295	0.1	[153]
SWNT	2.1	77	0.1	[153]
SWNT	8	80	12	[154]
SWNT	6.5	77	1.5	[155]
SWNT	0.05	296	3.6	[156]
SWNT	0.1	300∼	0.1	[157]
SWNT	1.6	R.T.		[158]
MWNT	3.4	290	10	[159]
MWNT	0.03	296	3.6	[156]
CNF	12.8	R.T.	11	[160]
CNF	ca. 5	300	10.1	[161]
CNF	1.2	296	10	[162]
CNF	0.7	R.T.	10.5	[163]
CNF	0.1	308	10	[164]
Nanostructured graphite	7.4	∼300	1.0	[165]

Note: R.T. is room temperature.

4.2.2. Catalysts

Application of porous carbon to catalyst is not a new and nontraditional example, but the surfaces of porous carbons still have unsolved problems particularly in the participation of surfaces to catalysis. So the topic concerning catalyst is also included in this section. A concise review is available for surveying the important research in the last two decades by Rodriguez-Reinoso [168].

The keywords to understanding the catalytic natures of carbon support are oxidation activity, electron transfer, high surface area, and spillover of hydrogen.

The oxidation ability of the carbon surface has been known and already used in flue gas desulfurization and denitrification reactions. Figure 35 shows a scheme for desulfurization on a carbon surface. The emitted SO_2 is adsorbed on the oxidation sites and then oxidized to SO_3 by the action of oxygen gas. The formed SO_3 is hydrated to sulfuric acid. Sulfuric acid first fills the micropores of porous carbon catalyst. After all of the micropores are filled with sulfuric acid, the reaction takes place on the macropore surface with lower reaction rate. Kisamori et al. claimed that the oxidation sites on the porous carbon are the surface defects induced by desorption of surface functional groups [169]. The desulfurization ability was not affected by surface areas of porous carbons, but porous carbons with higher nitrogen content showed excellent activity. The influence of surface treatments on the catalytic activity of carbon is an important topic in catalyst preparation [170–172]. Introduction of nitrogen onto the surfaces of porous carbons derived from a Victorian brown coal resulted in remarkable increases in the activity for both desulfurization and denitrification [173]. Kuhl et al. studied denitrification activity of sulfuric acid activated coke that was treated with ammonia [174]. They found that the denitrification activity increased when nitrogen was introduced in the range of 1–3%, and it saturated around 5%. However, the chemical structure and the mechanism of the introduced nitrogen are not clear.

Large surface areas of porous carbons have been used for catalyst supports. Porous carbon materials enable catalytic materials such as metals, metal oxides, and metal sulfides finely dispersed to exert high catalytic activity [168].

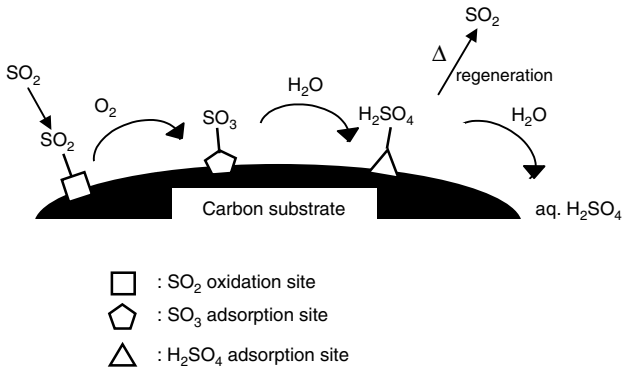


Figure 35. A schematic representation of desulfurization on a carbon surfaces.

Porous silica, alumina, and titania are more common catalyst supports, but the most striking difference of carbon supports from them is that we can anticipate electronic interaction between the supported metal and the carbon supports, because carbons are electronic conductors. An example can be seen in the Hoechst–Wakker reaction of ethylene to produce acetaldehyde. The catalyst, palladium chloride, is reduced when it is used for oxidizing ethylene and is regenerated by the presence of copper chloride. If palladium chloride is loaded onto a carbon support, the regenerating reagent is not required. This means that the surface of the carbon support promotes the reoxidation of palladium metal to form a continuous catalytic cycle. Similar electronic effects were also found for ammonia synthesis reactions.

Spillover is very interesting phenomenon involving platinum or nickel catalysts supported on carbon. Porous carbons do not adsorb hydrogen in the temperature range of room temperature to ca. 673 K, and uptake begins above 723 K gradually. On the other hand, the metal supported porous carbons adsorb hydrogen at temperatures as low as 473 K, and the adsorbed amount of hydrogen was 10^{21} atom/g-C, corresponding to the several tens or several hundreds times the amount of hydrogen which is expected from the metal loading [175]. This has been explained by the hydrogen transfer from the metal surfaces to the carbon surfaces as shown in Figure 36. The adsorbed hydrogen can easily be desorbed from the catalyst. In the figure dehydrogenation of hydrocarbon takes place on the carbon surface, and the removed hydrogen becomes spillover hydrogen. The produced olefin and hydrogen are effectively separated to inhibit recombination reaction.

4.2.3. Removing Environmentally Harmful Substances

Dioxins or environmental endocrine disruptors (environmental hormones) are emitted to environment which are too toxic and harmful and induce deformities or deaths, even with a very small quantity of these chemicals.

Dioxin is a collective name for chlorinated compounds including two benzene rings combined with ether linkage. The toxicity depends on the number and the substituted positions of chlorine on benzene rings. The optimum specification of the porous carbons for this purpose is as follows: pore size: 2–5 nm; specific surface area: >500 m^2/g; pore volume: >0.2 ml/g; particles size about 20 μm.

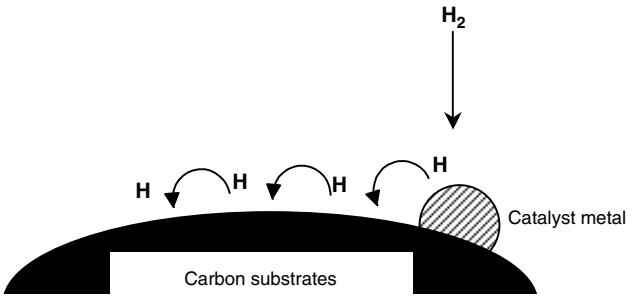


Figure 36. The concept of spillover of hydrogen in metal/carbon systems.

There are many chemical substances that are listed as environmental endocrine disruptors. Their removal by an adsorption technique is studied. Abe et al. studied the adsorption of 4-nonylphenol and bisphenol A on four types of activated carbons from different origins [176]. The porous carbons used there were particularly excellent adsorbents for bisphenol A by adsorbing up to saturation uptake even at concentrations as high as several mg/L.

In the practical situation for removing toxic substance from our environment, the concentration of the target molecules is very low as in the case of dioxin and furthermore the kinds of the suspicious molecules are various, which make the design of suitable porous carbon materials difficult. However, it is necessary to develop a design method for suitable porous carbon materials in order to protect our lives.

4.2.4. Electronics Uses

A battery is the most famous example of a carbon material for the electronic use. Graphite is used as an anode material in lithium ion batteries, and it shows a very close capacity to the theoretical capacity which is 372 mA hg^{-1} for this substance. On the other hand, some of the nongraphite carbons showed capacity exceeding the theoretical value. Tokumitsu et al. explained the high capacity on the basis of cavity model, in which lithium species are not only intercalated in carbon layers but also doped in cavities of which size was estimated as 0.5–1.5 nm by using a Fourier analysis of 002 diffraction profiles on an X-ray diffraction [177]. Although such carbons tend to show a large irreversible capacity, it is interesting to know the mechanism of lithium insertion and doping into these carbons from the viewpoint of structure of nongraphite carbons [178].

Carbons are also used in other battery applications, such as redox flow cells [179], zinc–air cells [180], etc. In polymer electrolyte membrane fuel cells (PEMFC), carbon materials are used in the gas diffusion electrode, bipolar plate, and the support for anode and cathode catalysts [124]. Three major barriers to technically overcome are developing bipolar plates, proton exchange membranes, and oxygen reduction catalysts. Because the working temperature for PEMFC is low as 353–373 K, the slow rate for oxygen reduction is the most serious problem. Usually porous carbon blacks are used as a catalyst support for platinum. In general carbon blacks with larger surface area can achieve a higher dispersion state of platinum, but we have to consider the pore structure of the carbon blacks, as it deeply affects the diffusion of oxygen gas and hydrogen ion. Sanyo Electric Co. Ltd. reported that carbon blacks with well-developed structure resulted in excellent oxygen diffusion. Matsushita Electric Co. Ltd. examined a number of carbon blacks and found that the carbon blacks that have a large amount of pores less than 10 nm are not suitable for the catalyst support. Carbon blacks for PEMFC applications are just for general use, so we need to develop a suitable carbon support. Iijima et al. recently found carbon nanohorns [20], which can be prepared just by laser ablation at room temperature. They showed a new type of carbon material especially for PEMFC application, as described in Section 2.2.

EDLC can be a complementary power source for secondary battery or internal combustion engines, because the

power density of EDLC is very high. This device is based on the electric capacity of the electrochemical double layer, which is formed at the interface of electrode and electrolyte. If the electrode is polarized positive, anions are attracted to the vicinity to the electrode surface. As a result opposite charges are aligned in face in a very short distance. This region is the so-called Helmholtz layer with a thickness of ca. 1 nm depending on the concentration of electrolyte. Hence the electric capacitance of the Helmholtz layer becomes considerably large. Porous carbon materials are commonly used for the electrodes of EDLC. As can be seen from the principle, the performance of the capacitor profoundly depends on the carbon pore structure. Yoshida et al. conducted a series of studies on the relation between the surface properties of the carbon electrode and the capacitance performances [181–183]. Their conclusion is as follows: increasing to pores larger than 2 nm leads to a good performance at lower temperature, and a smaller amount of surface functional groups is recommended for minimizing leak current. It is desired to prepare porous carbons with higher surface area and higher densities, but these requirements are oppositely demanding, so we must find an optimum pore structure for EDLC electrodes. Endo et al. also reported a morphological characterization and EDLC properties of activated carbons derived from a resin [184]. In this study they analyze the correlation between the capacitance and pore size distribution by a computer simulation. It would provide information on the relation between EDLC and ion-size related charge transfer solvation.

4.2.5. Electrochemical Application

Graphite is an anisotropic crystal, so that it has two types of surfaces. One is the basal plane normal to the stacking graphene sheets and another is the edge plane perpendicular to the basal plane. Kawakubo et al. demonstrated the difference in the activities between these crystal planes by using a redox reaction between ferricyanide and ferrocyanide [185]. Cyclic voltammograms of this study are presented in Figure 37. As can be seen in this figure, the edge plane of graphite is more active for electron transfer at the electrode/electrolyte interface than the basal plane of graphite. They also studied the electrochemical activity of the lead of mechanical pencils and found the anisotropic nature of the materials. Ozaki et al. prepared carbon materials by carbonizing poly(furfuryl alcohol), phenolic resin or a brown coal by mixing with iron compounds, and found that the obtained carbons showed electrocatalytic activity comparable to a platinum electrode [56–59]. They also found that the high electrochemical activity has a good relation between the formation of a special form of carbon, Ts carbon, of which the surface may be occupied by edge-plane-like structures.

Electrochemical separation or decomposition of impurities in water is another example of utilizing porous carbon materials in electrochemical applications. Capacitive deionization (CDI) of water is another application of the same principle as EDLC [113]. Figure 38 shows the principle of the CDI process for making fresh water from salty water. The striking difference of CDI from EDLC is that the electrolyte always flows through the cell. Electrodes for this application are usually made with ACFs, and sometimes with

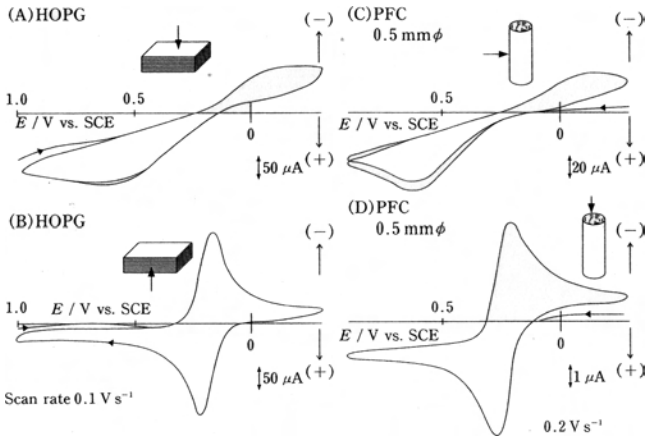


Figure 37. Cyclicvoltammograms for different carbon materials in a $Fe(CN)_6^{3-}$ aqueous solution. (A) and (B) Highly oriented pyrolytic carbons exposing different crystal surfaces. (C) and (D) Leads for mechanical pencils exposing different orientations. Reprinted with permission from [185], T. Kawakubo et al., *TANSO* 169, 201 (1995). © 1995, Carbon Society of Japan.

activated carbons with binders. Recent development of the sol–gel technique has enabled the preparation of carbon gels with desired shapes such as monoliths or films.

The flow-through-type electrochemical reactor is a very interesting application of porous carbon, although this is not so common in practical processes. Matsunaga et al. developed a flow-type reactor for electrochemical sterilization [186, 187]. This is based on their findings that applying 1.0 V of potential to the electrode could suppress the multiplication of bacteria on the electrode, and they could be taken off from the electrode by applying −0.8 V. By applying both voltages alternately, the multiplication of the bacteria was found to be considerably suppressed. They used ACF but changed to use titanium nitride as the porous electrode because of the lack of mechanical strength of the ACF.

There are several studies on electrochemical removal of pollutions from water or gases, such as dechlorination of organic compounds by using a flow-type reactor [188], electrochemical membrane separator of chlorine from gaseous hydrogen chloride wastes [189], and electrochemical

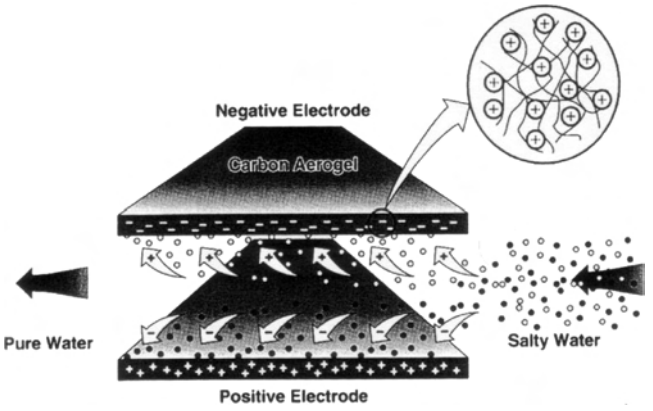


Figure 38. Concept of capacitive deionization of water with carbon electrodes. Reprinted with permission from [113], J. C. Farmer et al., *J. Electrochem. Soc.* 143, 159 (1996). © 1996, Electrochemical Society.

oxidation of bisphenol A [190]. These technologies are interesting and important not only from technical standpoints, but also an academic point of view of a coupling of catalysis and electrolysis occurring on carbon surfaces.

4.2.6. Medical Application

There are three main areas in the medical use of activated carbons: (1) as an oral adsorbent-antidote against acute poisoning; (2) for blood and other fluid purification using extracorporeal techniques; (3) as an odor adsorbing dressing for treatment of malodorous wounds [191]. The biocompatibility of the activated carbons with blood and so on is required for these uses. At the same time the presence of mesopores is of particular importance in order to remove large molecules such as proteins, circulating immune complexes, or endotoxin.

5. CONCLUSION

We reviewed the present status of the science and technology surrounding the nanoporous carbons. As a result it was recognized that the formation and the control of pore size ranging from micropore to mesopore are extremely important to meet the future demands for these carbons. While the improvement of the traditional preparation process is now being continued, new approaches to produce excellent porous carbons with outstanding properties, such as the template method, the polymer blend method, and the sol–gel process have been proposed and their performance of the products is being investigated by many researchers, as briefly described. In addition, the appearance of novel carbons such as carbon nanotubes brings about the stimulation and acceleration of the research activities in the carbon world, and at the same time the application field is expanded further. If the existing porous carbons such as activated carbons and ACFs would be examined thoroughly from the different points of view, they could play a more important role in the fields of environment, energy storage, catalysts, fuel cells, electronics, medical applications, and so forth.

GLOSSARY

BET surface area A popular surface area obtained from nitrogen adsorption isotherm, using a formula derived by Brunauer, Emmett, and Teller.

Carbon aerogel Mesoporous carbon spheres obtained by carbonization of organic gels which are dried at supercritical conditions.

Carbon fiber (PAN-based and pitch-based) Fibrous carbon materials obtained by carbonization of fibrous raw materials derived from PAN and pitch.

Electrochemical double layer capacitor (EDLC) A kind of capacitor utilizing the capacitance of electrochemical double layer formed near the electrode surface, showing extraordinarily high capacitance.

Mesophase pitch Optically anisotropic pitch due to liquid crystal nature.

Nongraphitizable and graphitizable carbons (hard and soft carbons) Type of carbon materials discriminated by the crystallinity after high temperature treatment (e.g., 3000 °C).

Polymer electrolytic membrane fuel cells (PEMFC) Type of fuel cell utilizing proton exchangeable polymer as an electrolyte of a fuel cell. It is expected to work at lower temperature than conventional fuel cells and to be utilized in domestic and on-board power generation in the near future.

sp, sp^2, and sp^3 Hybridizaton Electronic states of atoms determining molecular shapes. Carbyne, graphite, and diamond are available for carbon atom.

Supercritical fluid (SCF) A fluid at the state over its critical point, showing high solubility and low viscosity. It is also recognized as a new type of solvent for reactions.

REFERENCES

1. R. E. Franklin, *Proc. Roy. Soc. London Ser. A* 209, 196 (1951).
2. S. Mrozowski, in "Proc. of the Conference on Carbon," p. 31. Univ. of Buffalo, 1956.
3. G. M. Jenkins and K. Kawamura, *Nature* 231, 175 (1971).
4. M. Shiraishi, in "Introduction to Carbon Materials" (Carbon Society of Japan, Ed.), revised version, p. 29. Science and Technology Publishing, Tokyo, 1984. [in Japanese].
5. J. D. Brooks and G. H. Tayler, *Carbon* 3, 185 (1965).
6. F. Derbyshire, M. Jagtoyen, R. Andrews, A. Rao, I. Martin-Gullon, and E. A. Grulke, in "Chemistry and Physics of Carbon" (L. R. Radovic, Ed.), Vol. 27, p. 1. Dekker, New York, 2001.
7. L. R. Radovic and C. Sudhakar, in "Introduction to Carbon Technologies" (H. Marsh, E. A. Heintz, and F. Rodriguez-Reinoso, Eds.), p. 103. Publications de la Universidad de Alicante, 1997.
8. E. Frackwiak and F. Beguin, *Carbon* 39, 937 (2001).
9. K. S. W. Sing, D. H. Everett, R. A. W. Haul, L. Moscou, R. A. Pierotti, J. Rouqerol, and T. Siemieniewska, *Pure Appl. Chem.* 57, 603 (1985).
10. S. J. Gregg and K. S. W. Sing, in "Adsorption, Surface Area and Porosity," 2nd ed. Academic Press, London, 1982.
11. M. Toyoda and M. Inagaki, *Carbon* 38, 199 (2000).
12. T. L. Cook, C. Komodromos, D. F. Quinn, and S. Ragan, in "Carbon Materials for Advanced Technology" (T. D. Burchell, Ed.), p. 269. Pergamon, Elmsford, NY, 1999.
13. M. B. Rao and S. Sircar, *J. Membr. Sci.* 85, 253 (1993).
14. Y.-G. Seo, G.-H. Kum, and A. Seaton, *Fundam. Adsorption* 7, 385 (2001).
15. M. S. Strano and H. C. Foley, *Carbon* 40, 1029 (2002).
16. M. Inagaki and K. Kaneko, in "Commentary on Carbon Family–Variety and Evaluation" (M. Inagaki, Ed.), p. 156. Agune Syoufuu Sha, Tokyo, 2001. [in Japanese].
17. N. Yoshizawa, Y. Yamada, M. Shiraishi, K. Kaneko, and N. Setoyama, *J. Chem. Soc., Faraday Trans.* 92, 2297 (1996).
18. S. Iijima, *Nature* 354, 56 (1991).
19. B. K. Pradhan, G. U. Sumanasekera, C. K. W. Adu, H. Romero, and P. C. Eklund, *Mater. Res. Soc. Symp. Proc.* 633 (2001).
20. S. Iijima, M. Yudasaka, R. Yamada, S. Bandow, K. Suenaga, F. Kokai, and K. Takahashi, *Chem. Phys. Lett.* 309, 165 (1999).
21. K. Murata, K. Kaneko, F. Kokai, K. Takahashi, M. Yudasaka, and S. Iijima, *Chem. Phys. Lett.* 331, 14 (2000).
22. S. Brunauer, L. S. Deming, W. E. Deming, and E. Teller, *J. Am. Chem. Soc.* 62, 1723 (1940).
23. I. Langmuir, *J. Am. Chem. Soc.* 40, 1361 (1918).
24. S. Brunauer, P. E. Emmett, and E. Teller, *J. Am. Chem. Soc.* 60, 309 (1938).

25. M. M. Dubinin, in "Progress in Surface and Membrane Science" (D. A. Cadenhead, Ed.), p. 9. Academic Press, New York, 1975.
26. M. M. Dubinin, *Carbon* 23, 373 (1985).
27. B. C. Lippens and J. H. de Boer, *J. Catal.* 4, 319 (1965).
28. K. S. W. Sing, in "Surface Area Determination" (D. H. Everett and R. H. Ottewill, Eds.), p. 25. Butterworths, London, 1970.
29. S. J. Gregg and K. S. W. Sing, in "Adsorption, Surface Area and Porosity," 2nd ed. Academic Press, London, 1982.
30. E. P. Barrett, L. G. Joyner, and P. P. Halenda, *J. Am. Chem. Soc.* 73, 373 (1951).
31. R. W. Cranston and F. A. Inkley, *Adv. Catal.* 9, 143 (1957).
32. D. Dollimore and G. R. Heal, *J. Appl. Chem.* 14, 109 (1964).
33. D. Dollimore and G. R. Heal, *J. Colloid Interface Sci.* 33, 508 (1970).
34. O. Kratky, in "Small Angle X-ray Scattering" (O. Glatter and O. Kratky, Eds.), p. 3. Academic Press, London, 1982.
35. M. Ruile, T. Kasu, N. Setoyama, T. Suzuki, and K. Kaneko, *J. Phys. Chem.* 98, 9494 (1994).
36. K. Fukuyama, Y. Kasahara, N. Kasahara, A. Oya, and K. Nishikawa, *Carbon* 39, 287 (2001).
37. A. Guinier and G. Fournet, in "Small Angle Scattering of X-ray." Wiley, New York, 1955.
38. G. Porod, "Small Angle X-ray Scattering" (O. Glatter and O. Kratky, Eds.), p. 32. Academic Press, London, 1982.
39. B. MacEnaney and T. J. Mays, in "Characterization of Porous Solids III" (J. Rouquerol, F. Rodriguez-Reinoso, K. S. W. Sing, and K. K. Unger, Eds.), p. 327. Elsevier Science, New York, 1994.
40. K. Oshida, K. Kogiso, K. Matsubayashi, K. Takeuchi, S. Kobayashi, M. Endo, M. S. Dresselhaus, and G. Dresselhaus, *J. Mater. Res.* 10, 2507 (1995).
41. V. Vignal, A. W. Morawski, H. Konno, and M. Inagaki, *J. Mater. Res.* 14, 1102 (1999).
42. M. Inagaki, V. Vignal, A. H. Konno, and W. Morawski, *J. Mater. Res.* 14, 3152 (1999).
43. M. Inagaki, M. Sunahara, A. Shindo, V. Vignal, and H. Konno, *J. Mater. Res.* 14, 3208 (1999).
44. M. Inagaki and Y. Suwa, *Carbon* 39, 915 (2001).
45. K. Oshida, N. Ekinaga, M. Endo, and M. Inagaki, *TANSO* 173, 142 (1996) [in Japanese].
46. R. C. Bansal, J.-B. Donnet, and F. Stoeckli, in "Active Carbon," Chapter 3. Dekker, New York, 1988.
47. R. C. Bansal, J.-B. Donnet, and F. Stoeckli, in "Active Carbon," Chapter 1. Dekker, New York, 1988.
48. P. L. Walker, Jr., F. Rusinko, Jr., and L. G. Austin, *Adv. Catal.* 11, 136.
49. R. T. K. Baker, C. R. F. Lund, and J. J. Chludzinski, Jr., *J. Catal.* 87, 255 (1984).
50. C. A. Mims, J. J. Chludzinski, Jr., J. K. Pabst, and R. T. K. Baker, *J. Catal.* 88, 97 (1984).
51. A. Oya, S. Yoshida, J. Alcaniz-Monge, and A. Linares-Solano, *Carbon* 33, 1085 (1995).
52. J. Ozaki and A. Oya, *HYOMEN* (*Surface*) 35, 535 (1997) [in Japanese].
53. H. Tamai, T. Kakii, Y. Hirota, T. Kumamoto, and H. Yasuda, *Chem. Mater.* 8, 456 (1996).
54. H. Tamai, M. Ikeuchi, S. Kojima, and H. Yasuda, *Adv. Mater.* 9, 55 (1997).
55. H. Nakagawa, K. Watanabe, Y. Harada, and K. Miura, *Carbon* 37, 1455 (1999).
56. J. Ozaki, M. Mitsui, and Y. Nishiyama, *Carbon* 36, 131 (1998).
57. J. Ozaki, M. Mitsui, Y. Nishiyama, J. D. Cashion, and L. J. Brown, *Chem. Mater.* 10, 3386 (1998).
58. J. Ozaki, K. Nozawa, and A. Oya, *Chem. Lett.* 1998, 573.
59. J. Ozaki, K Nozawa, L. J Brown, J. D. Cashion, and A. Oya, submitted for publication.
60. O. Kadlec, A. Varhanikova, and A. Zukal, *Carbon* 8, 321 (1970).
61. A. N. Weunerberg and T. M. O'Grady, U.S. Patent 4, 082, 694, 1978.
62. H. N. S. Schafer, U.S. Patent 4, 039, 473, 1977.
63. R. Ryoo, S. H. Joo, M. Kruk, and M. Jaroniec, *Adv. Mater.* 13, 677 (2001).
64. M. Morina-Sabio, M. T. Gonzalez, F. Rodriguez-Reinoso, and A. Sepulveda-Escribano, *Carbon* 34, 505 (1996).
65. F. Rodriguez-Reinoso, A. C. Pastor, H. Marsh, and M. A. Martinez, *Carbon* 38, 379 (2000).
66. S.-J. Park and K.-D. Kim, *Carbon* 39, 1741 (2001).
67. P. J. M. Carrott, J. M. V. Nabais, M. M. L. Ribeiro-Carrott, and J. A. Pajares, *Carbon* 39, 1543 (2001).
68. T. Kyotani, *Carbon* 38, 269 (2000).
69. J. H. Knox and M. T. Gilbert, U.S. Patent 4, 263, 268, 1981.
70. M. T. Gilbert, J. H. Knox, and B. Kaur, *Chromatographia* 16, 138 (1982).
71. J. H. Knox, B. Kaur, and G. R. Millward, *J. Chromatogr.* 352, 3 (1986).
72. K. Kamegawa and H. Yoshida, *Carbon* 35, 631 (1997).
73. K. Kamegawa and H. Yoshida, *J. Mater. Sci.* 34, 3105 (1999).
74. T. Hyeon, S. Han, J. Lee, and K. Sohn, *Fuel. Chem. Div. Prep.* 46, 268 (2001).
75. Z. Li and M. Jaroniec, *J. Am. Chem. Soc.* 123, 9208 (2001).
76. D. Kawashima, T. Aihara, Y. Kobayashi, T. Kyotani, and A. Tomita, *Chem. Mater.* 12, 3397 (2000).
77. T. Kyotani, N. Sonobe, and A. Tomita, *Nature* 331, 331 (1988).
78. N. Sonobe, T. Kyotani, and A. Tomita, *Carbon* 29, 61 (1991).
79. T. Kyotani, T. Mori, and A. Tomita, *Chem. Mater.* 6, 2138 (1994).
80. T. Kyotani, T. Nagai, S. Inoue, and A. Tomita, *Chem. Mater.* 9, 609 (1997).
81. Z. Ma, T. Kyotani, and A. Tonita, *Chem. Commun.* 2365 (2000).
82. J. Lee, S. Yoon, T. Hyeon, S. M. Oh, and K. B. Kim, *Chem. Commun.* 2177 (1999).
83. R. Ryoo, S. H. Joo, and S. Jun, *J. Phys. Chem. B* 103, 7743 (1999).
84. S. H. Joo, S. J. Choi, I. Oh, J. Kwak, Z. Liu, O. Terasaki, and R. Ryoo, *Nature* 412, 169 (2001).
85. F. P. Dousek and J. Jansta, *Carbon* 18, 13 (1980).
86. L. Kavan, *Chem. Rev.* 97, 3061 (1997).
87. S. Shiraishi, H. Kurihara, H. Tsubota, A. Oya, Y. Soneda, and Y. Yamada, *Electrochem. Solid-State Lett.* 4, A5 (2001).
88. T.-T. Liang, Y. Yamada, N. Yoshizawa, S. Shiraishi, and A. Oya, *Chem. Mater.* 13, 2933 (2001).
89. V. B. Fenelonov, V. A. Likholobov, A. Yu. Derevyankin, and M. S. Melgunov, *Catal. Today* 42, 341 (1998).
90. A. A. Zakhilov, R. H. Baughman, Z. Iqbal, C. Cui, I. Khayrullin, S. Socrates, O. Dantas, J. Marti, and V. G. Ralchenko, *Science* 282, 894 (1998).
91. J. Ozaki, N. Endo, W. Ohizumi, K. Igarashi, M. Nakahara, and A. Oya, *Carbon* 35, 1031 (1997).
92. H. Hatori, Y. Yamada, and M. Shiraishi, *Carbon* 30, 303 (1992).
93. H. Hatori, Y. Yamada, and M. Shiraishi, *J. Appl. Polym. Sci.* 57, 871 (1995).
94. T. Takeichi, Y. Yamazaki, M. Zuo, A. Ito, A. Matsumoto, and M. Inagaki, *Carbon* 39, 257 (2001).
95. J. Ozaki, W. Ohizumi, and A. Oya, *Carbon* 38, 1515 (2000).
96. J.-B. Yang, L.-C. Ling, L. Liu, F.-Y. Kang, Z.-H. Huang, and H. Wu, *Carbon* 40, 911 (2002).
97. J. Ozaki, W. Ohizumi, N. Endo, A. Oya, S. Yoshida, T. Iizuka, M. C. Roman-Martinez, and A. Linares-Solano, *Carbon* 35, 1076 (1997).
98. J. Ozaki, W. Ohizumi, A. Oya, M. J. Illan-Gomez, M. C. Roman-Martinez, and A. Linares-Solano, *Carbon* 38, 778 (2000).
99. K. Takakura, H. Tanimoto, and M. Kijima, in "Proceedings of the 28th Annual Meeting of the Carbon Society of Japan," Kiryu, p. 208. 2001 [in Japanese].

100. K. Lenghaus, G. G. Qiao, D. H. Solomon, C. Gomez, F. Rodriguez-Reinoso, and A. Sepulveda-Escribano, *Carbon* 40, 743 (2002).
101. R. W. Pekala, *J. Mater. Sci.* 24, 3221 (1984).
102. R. W. Pekala, C. T. Alviso, F. M. Kong, and S. S. Hulsey, *J. Non-Cryst. Solids* 145, 90 (1992).
103. H. Tamon, H. Ishizaka, M. Mikami, and M. Okazaki, *Carbon* 35, 791 (1997).
104. H. Tamon, H. Ishizaka, T. Yamamoto, and T. Suzuki, *Carbon* 37, 2049 (1999).
105. Y. Hanzawa, H. Hatori, N. Yoshizawa, and Y. Yamada, *Carbon* 40, 575 (2002).
106. R. W. Pekala, C. T. Alviso, X. Lu, J. Gross, and J. Fricke, *J. Non-Cryst. Solids* 188, 34 (1995).
107. G. C. Ruben and R. W. Pekala, *J. Non-Cryst. Solids* 186, 219 (1995).
108. G. Biesmans, A. Martens, T. Woignier, and J. Phalippou, *J. Non-Cryst. Solids* 225, 64 (1998).
109. W.-C. Li, A.-H. Lu, and S.-C. Guo, *Carbon* 39, 1989 (2001).
110. R. W. Pekala, J. C. Farmer, C. T. Alviso, T. D. Tran, S. T. Mayer, J. M. Miller, and B. Dunn, *J. Non-Cryst. Solids* 225, 74 (1998).
111. S. T. Mayer, R. W. Pekala, and J. L. Kaschmitter, *J. Electrochem. Soc.* 140, 446 (1993).
112. H. Proebstle, C. Schmitt, and J. Fricke, *J. Power Sources* 105, 189 (2002).
113. J. C. Farmer, D. V. Fix, G. V. Mack, R. W. Pekala, and J. F. Poco, *J. Electrochem. Soc.* 143, 159 (1996).
114. M Glora, M. Wiener, H. Proebstle, G. Reichnauer, and J. Fricke, in "Extended Abstracts, An International Conference on Carbon (Carbon'02)," Beijing, 2002.
115. J. Wang, L. Angnes, H. Tobias, R. A. Roesner, K. C. Hong, R. S. Glass, F. M. Kong, and R. W. Pekala, *Anal. Chem.* 65, 2300 (1993).
116. C. U. Pittman, Jr., W. Jiang, Z. R. Yue, and C. A. Leon y Leon, *Carbon* 37, 85 (1999).
117. Z. R. Yue, W. Jiang, L. Wang, H. Toghiani, S. D. Gardner, and C. U. Pittman, *Carbon* 37, 1607 (1999).
118. C. U. Pittman, Jr., W. Jiang, Z. R. Yue, S. Gardner, L. Wang, H. Toghiani, and C. A. Leon y Leon, *Carbon* 37, 1797 (1999).
119. T. Sako, T. Sugeta, K. Otake, N. Nakazawa, M. Sato, and M. Tsugumi, *J. Chem. Eng. Jpn.* 30, 342 (1997).
120. J. Ozaki, S. K. I. Djaja, and A. Oya, *Ind. Eng. Chem. Res.* 39, 245 (2000).
121. E. Reverchon, G. Dell Porta, A. D. Trolio, P. Pallado, and A. Stassi, *Ind. Eng. Chem. Res.* 34, 4087 (1995).
122. E. Reverchon, *J. Supercrit. Fluids* 15, 1 (1999).
123. Y.-Y. Li, K. Mochidzuki, A. Sakoda, and M. Suzuki, *Carbon* 39, 2143 (2001).
124. J. Guo and A. C. Lau, *Carbon* 38, 1985 (2000).
125. W. Lu and D. D. L. Chung, *Carbon* 39, 39 (2001).
126. M. Inagaki and L. R. Radovic, *Carbon* 40, 2263 (2002).
127. A. C. Dillon, K. M. Jones, T. A. Bekkedahl, C. H. Kiang, D. S. Bethune, and M. J. Heben, *Nature* 386, 377 (1997).
128. M. S. Dresselhaus, K. A. Williams, and P. C. Eklund, *Mater. Res. Bull.* 24, 45 (1999).
129. M. Eswaramoorthy, R. Sen, and C. N. R. Rao, *Chem. Phys. Lett.* 304, 207 (1999).
130. S. Inoue, N. Ichikuni, T. Suzuki, T. Uematsu, and K. Kaneko, *J. Phys. Chem.* 102, 4689 (1998).
131. J. N. Barisci, G. G. Wallace, and R. H. Baughman, *J. Electroanal. Chem.* 488, 92 (2000).
132. R. C. Bansal, J.-B. Donnet, and F. Stoeckli, in "Active Carbon." Dekker, New York, 1988.
133. T. L. Cook, C. Komodromos, D. F. Quinn, and S. Fagan, in "Carbon Materials for Advanced Technologies" (T. D. Burchell, Ed.), Ch. 9, p. 269. Pergamon, Amsterdam, 1999.
134. T. Masuda, E. Isobe, and T. Higashiyama, *J. Am. Chem. Soc.* 105, 7473 (1983).
135. K. Takada, H. Matsuya, T. Masuda, and T. Higashiyama, *J. Appl. Polym. Sci.* 30, 1605 (1985).
136. Y. P. Yampolskii, V. P. Shantorovich, F. P. Chernyakovskii, A. I. Kornilov, and N. A. Plate, *J. Appl. Polym. Sci.* 47, 85 (1993).
137. V. Bondar, A. Alentiev, T. Masuda, and U. Yampolskii, *Macromol. Chem. Phys.* 198, 1701 (1997).
138. A. Golovoy, in "Proceedings of the 1983 Compressed Natural Gas Conference, " p. 129. Society of Automotive Engineers, Pittsburgh, 1983.
139. H. D. Gesser, A. Rocbon, A. E. Lemire, K. J. Masters, and M. Raudsepp, *Zeolites* 4, 22 (1984).
140. R. Cracknell, P. Gordon, and K. Gubbins, *J. Phys. Chem.* 97, 494 (1993).
141. N. Parkyns and D. Quinn, in "Porosity in Carbons" (J. W. Patrick and E. Arnold, Eds.), Ch. 11, p. 300. London, 1995.
142. T. A. Brady, M. R. Ostam-Abadi, and M. J. Rood, *Gas Sep. Purification* 10, 97 (1996).
143. T. Tan and K. E. Gubbins, *J. Phys. Chem.* 94, 6061 (1992).
144. K. R. Matranga, A. L. Myers, and E. D. Glandt, *Chem. Eng. Sci.* 47, 1569 (1991).
145. J. Alcaniz, M. A. de la Cassa-Lillo, D. Cazoria-Amoros, and A. Linares-Solano, *Carbon* 35, 291 (1997).
146. J. Sun, M. J. Rood, M. Rostam-Abadi, and A. A. Lizzo, *Gas Sep. Purification* 10, 91 (1996).
147. F. Catula, M. Molina-Sabio, and F. Rodriguez-Reinoso, *Carbon* 29, 999 (1991).
148. J. A. F. MacDonald and D. F. Quinn, *Carbon* 34, 1103 (1996).
149. E. A. Muller and K. E. Gubbins, *Carbon* 36, 1433 (1998).
150. H. M. Cheng, Q. H. Yang, and C. Liu, *Carbon* 39, 1447 (2001).
151. A. C. Dillon, K. E. Gilbert, J. L. Alleman, T. Gennett, K. M. Jones, P. A. Parilla, and M. J. Heben, in "Proc. of the 2001 DOE Hydrogen Program Review," NREL/CP-570-30535.
152. C. Liu, Y. Y. Fan, M. Liu, H. T. Cong, H. M. Cheng, and M. S. Dresselhaus, *Science* 286, 1127 (1999).
153. N. Nishimiya, K. Ishigaki, H. Takikawa, M. Ikeda, Y. Hibi, T. Sakakibara, A. Matsumoto, and K. Tsustumi, *J. Alloys Compd.* 339, 275 (2002).
154. Y. Ye, C. C. Ahn, C. Witham, B. Fultz, J. Liu, A. G. Rinzler, D. Colbert, K. A. Smith, and R. E. Smalley, *Appl. Phys. Lett.* 74, 2307 (1999).
155. B. K. Pradhan, G. U. Sumanasekera, C. K. W. Adu, H. E. Romero, K. A. Williams, and P. C. Eklund, *Physica B* 321 (2002).
156. G. G. Tibbetts, G. P. Meisner, and C. H. Olk, *Carbon* 39, 2291 (2001).
157. M. Hirscher, M. Becher, A. Quintel, V. Skakalova, Y. M. Choi, S. Roth, I. Stepanek, P. Bernier, A. Leonhardt, and J. Fink, in "Extended Abstracts of Eurocarbon" p. 91. 2000.
158. A. K. M. Fazle Kibria, Y. H. Mo, K. S. Park, K. S. Nahm, and M. H. Yun, *Int. J. Hydrogen Energy* 26, 823 (2001).
159. H. Zhu, A. Cao, X. Li, C. Xu, Z. Mao, D. Ruan, J. Liang, and D. Wu, *Appl. Surf. Sci.* 178, 50 (2001).
160. Y. Y. Fan, B. Liao, M. Liu, Y. L. Wei, M. Q. Lu, and H. M. Cheng, *Carbon* 37, 1649 (1999).
161. H. M. Cheng, C. Liu, Y. Y. Fan, F. Li, G. Su, H. T. Cong, L. L. He, and M. Liu, *Z. Metallkd.* 91, 306 (2000).
162. R. Ströbel, L. Jörissen, T. Schliermann, V. Trapp, W. Schütz, K. Bohmhammel, G. Wolf, and J. Garche, *J. Power Sources* 84, 221 (1999).
163. E. Poirier, R. Chahine, and T. K. Bose, *Int. J. Hydrogen Energy* 26, 831 (2001).
164. T. Kiyobayashi, H. T. Takeshita, H. Tanaka, N. Takeichi, A. Züttel, L. Schlapbach, and N. Kuriyama, *J. Alloys Compd.* 330–332, 666 (2002).

165. S. Orimo, T. Matsushima, H. Fujii, T. Fukunaga, and G. Majer, *J. Appl. Phys.* 90, 1545 (2001).
166. P. David, T. Piquero, K. Metenier, Y. Pierre, J. Demoment, and A. Lecas-Hardit, in "Extended Abstracts of International Conference on Carbon (Carbon '02)," Beijing, 2002, G044.
167. H. Takagi, H. Hatori, Y. Soneda, and Y. Yamada, *TANSO* 205, 231 (2002) [in Japanese].
168. F. Rodriguez-Reinoso, *Carbon* 36, 159 (1998).
169. S. Kisamori, K. Kuroda, S. Kawano, I. Mochida, Y. Matsumura, and M. Yoshikawa, *Energy Fuels* 8, 1337 (1994).
170. S. Wang and G. Q. Lu, *Carbon* 36, 283 (1998).
171. Z. H. Zhu, L. R. Radovic, and G. Q. Lu, *Carbon* 38, 451 (2000).
172. M. Gurrath, T. Kuretzky, H. P. Boehm, L. B. Okhlopkova, A. S. Lisitsyn, and V. A. Likholobov, *Carbon* 38, 1241 (2000).
173. H. Fujitsu, I. Mochida, G. J. Perry, and D. A. Allardice, *Fuel* 72, 109 (1993).
174. H. Kuhl, E. Richter, K. Knoblauch, and H. Huntgen, in "Reprints Carbon '86," p. 351, 1986.
175. A. J. Robell, E. V. Bellou, and M. Boudar, *J. Phys. Chem.* 68, 2748 (1964).
176. I. Abe, S. Iwasaki, J. Fukuhara, S. Nakanishi, and N. Kawasaki, *TANSO* 184, 234 (1998) [in Japanese].
177. K. Tokumitsu, H. Fujimoto, A. Mabuchi, and T. Kasuh, *Carbon* 37, 1599 (1999).
178. J. Gong and H. Wu, *Electrochim. Acta* 45, 1753 (2000).
179. N. H. Hagedorn, in "NASA Redox Storage System Development Project Final Report," DOE/NASA/12726-24, NASA, TM-83677, 1984.
180. B. C. H. Steele and A. Heinzel, *Nature* 414, 345 (2001).
181. A. Yoshida, I. Tanahashi, and A. Nishino, *IEEE Trans.* CHMT-10, 100 (1987).
182. A. Yoshida, I. Tanahashi, and A. Nishino, *Carbon* 28, 611 (1990).
183. A. Yoshida, I. Aoki, S. Nonaka, and A. Nishino, *J. Power Source* 60, 213 (1996).
184. M. Endo, Y. J. Kim, H. Ohta, K. Ishii, T. Inoue, T. Hayashi, Y. Nishimura, T. Maeda, and M. S. Dresselhaus, *Carbon* 40, 2613 (2002).
185. T. Kawakubo, Y. Suda, A. Negishi, and H. Kaneko, *TANSO* 169, 201 (1995).
186. T. Matsunaga, S. Nakasone, Y. Kitajima, and K. Horiguchi, *Biotechnol. Bioeng.* 43, 429 (1994).
187. M. Okochi, T. K. Lim, N. Nakamura, and T. Matsunaga, *Appl. Microbiol. Biotechnol.* 47, 1822 (1997).
188. S. M Kulikov, V. P. Plekhanov, A. I. Tsyganok, C. Schlimm, and E. Heitz, *Electrochim. Acta* 41, 527 (1996).
189. J. Johnson and J. Winnick, *Separation Purif. Technol.* 15, 223 (1999).
190. H. Kuramitz, Y. Nakata, M. Kawasaki, and S. Tanaka, *Chemosphere* 45, 37 (2001).
191. S. V. Mikhalovsky, in "Extended Abstracts, An International Conference on Carbon (Carbon '02)," Beijing, 2002.

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanoporous Materials

Dongyuan Zhao, Chengzhong Yu, Haifeng Yang

Fudan University, Shanghai 200433, China

CONTENTS

1. INTRODUCTION

Porous materials are a family of excellent adsorbents, catalysts, supporting materials, ion-exchanging agents, and micro-reactors. Because of their open-framework structures and large surface areas (including internal and external surface areas), these porous materials have been widely used in the petroleum and chemical industries and in air separation [1, 2]. According to the definition of the pore size from the International Union of Pure and Applied Chemistry (IUPAC) [3], porous materials can be classified into three groups: microporous materials (<2 nm), mesoporous materials (2–50 nm), and macroporous materials (>50 nm). These three classes of materials can be included in the so-called nanoporous materials (1–100 nm), which are found to be more and more important in the applications of large molecular conversion, optic-electric devices, and other high-tech industries [4, 5].

Nanoporous materials can be amorphous, quasi-crystalline, or crystalline. Methods such as X-ray diffraction (XRD) and electron diffraction (ED) can be used to characterize these materials [1, 2]. The pore structure is one of the most important parameters for nanoporous materials and can generally be characterized by use of physical adsorption of the gas (such as N_2, Ar, organic molecules, etc.) or by a liquid pressure method (mercury extrusion). Amorphous and quasi-crystalline porous materials such as silica gels, aluminum oxides, and activated carbons have been widely used in industry as catalysts or the supports for catalysts. However, because these materials possess no long-range ordered structural regularity and they have disordered pore structures and widely distributed pore sizes, zeolite-type microporous materials have long been used as substitutes.

Nanoporous materials with long-range ordered structures are discussed in this chapter, and the focus is on the newly developed classes of mesoporous materials. Ordered macroporous materials developed recently will also be involved. At the atomic level, most mesoporous and macroporous materials are amorphous, but they have properties similar to those of crystalline materials because their uniform pore structures possess long-range order at the meso- or macroscales, which can be reflected by well-defined XRD and ED patterns.

2. MESOPOROUS MATERIALS

Some disordered mesopores can be generated during a dealumination process of zeolite molecular sieves, but the pore sizes and the quantity of mesopores are much dependent on such a dealumination process, and somehow it is uncontrollable [1, 2]. Some layered materials such as natural clays and metal phosphates can be used to prepare mesoporous materials with the use of large inorganic precursors such as polycations (e.g., Al_{13}, $[Al_{13}O_4(OH)_{24}(H_2O)_{12}]^{7+}$) or orthosilicates (referred to as pillars) to support the lamellar structures [6–10]. The final products are called pillared clays or pillared molecular sieves [10]. These lamellar structures include smectite, bentonite, montmorillonites, and

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (263–285)

zirconium phosphate, etc. The pillared clays were once important candidates for catalysis in heavy oil cracking at the end of the 1970s. However, later experiments revealed that the surface acidity of such materials was quite weak, and the catalytic activities were greatly decreased because of the carbon coke; moreover, the hydrothermal stability of such pillared clays was not good enough for use in the cracking process of the petroleum industry [11]. Although natural clays and metal phosphates are crystalline, the pillars within these layered materials are not uniformly distributed; thus the resultant mesopores are not ordered [5, 11]. In some cases, the pore size distribution of aluminosilicates synthesized by the sol–gel method under strict conditions is quite narrow, but the mesopores are still randomly packed.

In 1992, Mobil researchers reported the synthesis of ordered mesoporous aluminosilicates (denoted M41S) with the use of long-chain cationic quaternary ammonium surfactants as structure-directing agents [12, 13]. In fact, ordered mesoporous materials were already reported early in 1971 [14]. A group of French scientists utilized such long-chain quaternary ammonium cationic surfactants to adjust the surface area and pore diameter of silica gels. According to their patent [15], the resulting silica gel had a two-dimensional (2D) hexagonal structure similar to that of mesoporous silica MCM-41 reported by the Mobil Company [15]. Unfortunately, this patent has been ignored because no XRD or transmission electron microscopy (TEM) evidence was provided [15]. A Japanese group began to synthesize mesoporous materials before 1990 [16]; the cationic surfactants were also employed to synthesize pillared lamellar clay, kanemite. It was found that when a high concentration of NaOH was used, the structure of kanemite was destroyed; instead, a new type of ordered mesoporous material was obtained, which was known as FSM-16 in later literatures [17]. Since the TEM images provided and the mechanism were not fully understood in the original report [16], this contribution did not draw enough attention from the international scientific involved in research on zeolite molecular sieves community.

On the contrary, after the successful synthesis of lamellar MCM-22 molecular sieves [18–20], Mobil scientists continued to study the pillared structures of MCM-22. It was also found that at the beginning stage the lamellar structure of MCM-22 was destroyed, and afterward new ordered mesostructures were formed. Detailed and systematic work had been carried out after the initial observation; it was found that these mesostructures had structures similar to those of liquid crystals formed from the employed surfactant. Finally, a new family of mesoporous materials (M41S) was discovered, including MCM-41, with a 2D hexagonal honeycomb structure (*p6mm*); MCM-48, with a 3D bicontinuous cubic structure (*Ia3d*); and MCM-50, with a stable pillared lamellar structure [12, 13].

The structures of mesoporous silica M41S are schematically shown in Figure 1. It should be noted that MCM-50 materials are stable mesoporous silica materials and have a pillared lamellar structure (pillared by silicate agglomerate species hydrolyzed from tetraethoxysilane, TEOS) instead of an unstable, simple lamellar structure [21, 22]. Since the first report by the Mobil Company, mesoporous materials have drawn tremendous attention internationally from

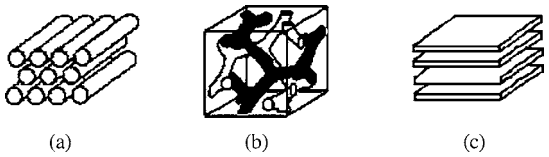


Figure 1. Schematic drawing of the pore structures of (a) MCM-41; (b) MCM-48; and (c) MCM-50.

researchers working on zeolite molecular sieves and in the materials research field. The success in the synthesis of mesoporous materials such as M41S is comparable to another great achievement with ZSM-5 zeolites in the 1970s, also by the Mobil Company [23]. Both materials have been synthesized by choosing organic structure-directing agents to control the pore size and shape of the resultant porous materials. The liquid-crystal-templating (LCT) mechanism in the case of mesoporous materials brings forward a true templating concept in the synthesis of porous materials; moreover, the restriction of the micropore (<1.2 nm) in the case of zeolite-type materials is no longer a problem [12, 13, 21, 22]: mesoporous materials with a uniform and controllable pore size between 1.5 and 40 nm offer new opportunities in many research fields in which large nanopore materials are essential.

The surfactants used in the synthesis of M41S materials were cationic quaternary ammonium salts with a hydrophilic headgroup and a hydrophobic long alkyl chain. In contrast to the organic templates in the synthesis of zeolite materials, the surfactants can generate complex supramolecular structures in aqueous solutions. At relatively low concentrations (>critical micelle concentration, CMC), the micelles can be formed, whereas at relatively high concentrations, liquid crystal phases are assembled. The surfactants and precursors of inorganic species cooperatively self-assemble into ordered composite mesostructures. This self-assembly process is similar to the formation of liquid crystals in solution. After the removal of surfactants, mesoporous materials are generated. The pore structure of mesoporous materials possesses long-range order, as reflected by well-defined XRD patterns at small angles ($2\theta = 2$–$8°$) (Fig. 2) [24], but the inorganic species (or inorganic wall) are amorphous; for example, a broad peak at $2\theta = 20$–$30°$ observed for mesoporous silica materials suggests that the silica walls are not crystalline [25–29].

3. FORMATION MECHANISM

In their early studies, Mobil scientists noticed that the formation of mesoporous materials is closely related to the assembly of traditional liquid crystals and biomineralization [12, 13]. The surfactants employed in the synthesis of mesoporous materials usually possess one or more polar headgroups and one long hydrophobic tail, whereas the lipid generally studied in life science has two hydrophobic tails. Traditional liquid crystals usually consist of one tail with a specific group (usually benzene groups). To explain the formation mechanism of MCM-41, the Mobil researchers first proposed a liquid-crystal-templating (LCT) mechanism [12, 13]. Based on high-resolution transmission electron

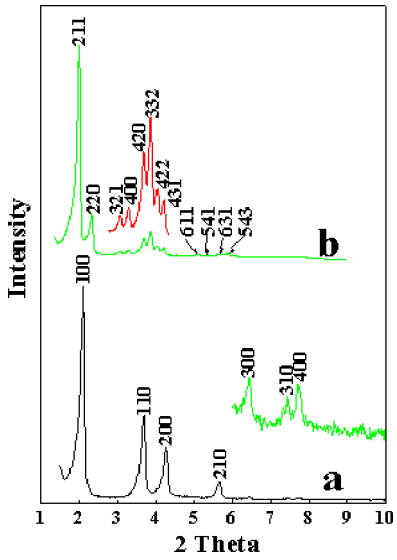


Figure 2. Representative XRD patterns of mesoporous materials. (a) Highly ordered MCM-41 prepared with cationic surfactant under basic conditions. (b) MCM-48 prepared with Gemini surfactant $C_{22\text{-}12\text{-}22}$ as a template at 100 °C.

microscopy (HRTEM) and XRD studies of MCM-41 materials [12, 13], which obtained results similar to those for liquid crystals formed in aqueous solutions, they proposed that the liquid crystal phase formed by the surfactants in aqueous solution was a template for the formation of an MCM-41 mesostructure. Mobil researchers suggested that the formation of liquid crystal phases of surfactants may occur (1) before and (2) after the addition of inorganic precursors [12, 13]. In this context, the LCT mechanism is "errorless," since there will be no other choices considering the formation of the liquid crystal phase.

When pathway (1) was utilized to explain the formation of mesoporous materials in aqueous solutions [12, 13], some conflicts occurred. In aqueous solutions, liquid crystal phases can only be observed at very high surfactant concentration. In the case of cetyltrimethyl ammonium bromide (CTAB), the hexagonal liquid crystal phase is formed when the concentration of CTAB is higher than 28 wt%, and the cubic phase occurs when the concentration of CTAB is higher than 80 wt%. On the other hand, MCM-41 can be synthesized at a very low CTAB concentration (~2 wt%) [24–28]. Even for the synthesis of cubic MCM-48 materials, the concentration of CTAB in starting compositions can be as low as 10 wt%. Later, Attard et al. prepared a liquid crystal phase by using surfactants and/or inorganic precursors under a nonaqueous phase [30]. With the use of this liquid crystal phase as a template, mesoporous materials were obtained. This observation supported the proposed pathway (1) by Mobil scientists; however, the formation of the liquid crystal phase prior to synthesis of mesoporous materials is somewhat difficult, and the synthesis of mesoporous materials strictly according to liquid crystal templating is not common in the literature [30].

Several formation mechanisms have been proposed to describe in more detail the synthesis pathway (2) where the liquid crystal phases are formed after the inorganic species were added. M. E. Davis et al. [25] suggested that surfactant molecules self-assembled into micelle rods first in aqueous solutions; these micelle rods could interact with inorganic silicate precursors and resulted in rod-like micelles of organic-inorganic hybrid composites that contained two or three layers of silicate oligomers (Fig. 3). These coated micelles agglomerated spontaneously to give a long-range ordered hexagonal mesostructure. After further condensation of silicates, the MCM-41 mesophase was obtained.

G. D. Stucky et al. proposed that the cooperative self-assembly of organic and inorganic species gave long-range ordered mesostructures [24, 27, 29]. Polymeric silicate anions interacted with the headgroups of cationic surfactants (Fig. 4) at the organic-inorganic interface; the charge density of the inorganic layer changed with the condensation of silicates and resulted in close packing of the long hydrophobic chain of the surfactants. The charge density matching between organic and inorganic species controlled the array of surfactant-inorganic hybrid materials. The pre-arrangement of organic surfactant molecules was not necessary in this mechanism. With the condensation of inorganic species, the charge density of the inorganic layer changed and the solid phase of the organic/inorganic composites was transferred. The final structure of the composite phase was governed by the extent of the reaction (the condensation of inorganic species) and by the charge density matching of surfactant/inorganic assemblies (Fig. 4).

Some questions may be argued with respect to Davis's mechanism. The length of the micelle rods is not uniform in solution; furthermore, long micelle rods must overcome a large energy gap to assemble into ordered organic-inorganic hybrid mesostructures. These facts contradict many observations; for example, the mesochannel of MCM-41 synthesized in the solution can be as long as several micrometers [31, 32]. In fact, at the surfactant concentration used to synthesize the mesoporous materials, the global micelles as well as rod-like micelles can be observed in solution; in this case, the composite phases other than hexagonal structures should coexist in the final products. Furthermore, this mechanism cannot be used to interpret the formation of cubic MCM-48 and lamellar mesostructured materials [24, 26]. Obviously, the limitation of Davis's mechanism restricted their general use in the synthesis of mesoporous materials. However, during the synthesis of mesostructured silica or germanium oxides, Yuan and Zhou [33] and Zhao et al. [34] observed that single rod-like inorganic or organic/inorganic materials could be obtained, suggesting that in some cases Davis's mechanism might be applicable.

The cooperative self-assembly mechanism proposed by G. D. Stucky et al. is quite general and thus has been accepted by most scientists. This mechanism can be widely used in different synthesis systems, and the general principle can be applied to the design of new materials.

Figure 3. Self-assembly of silicate-encapsulated micelle rods into hexagonal mesophase. Reprinted with permission from [4], J. Y. Ying et al., *Angew. Chem. Int. Ed. Engl.* 38, 56 (1999). © 1999, Wiley.

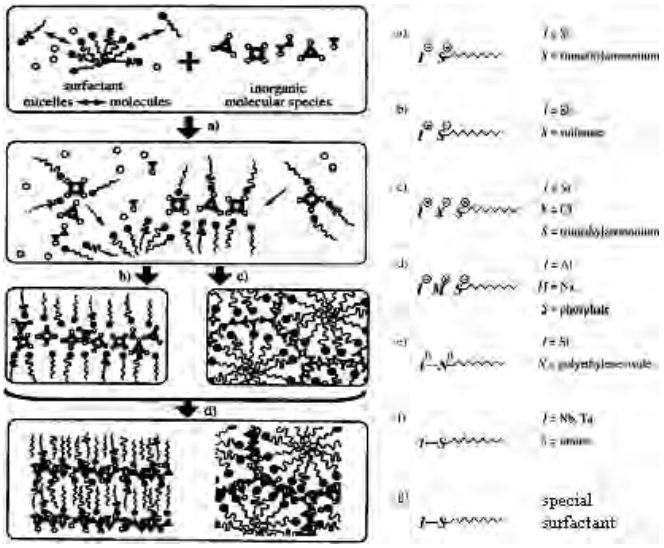


Figure 4. Cooperative templating model (left) and the synthesis approach (right) for the synthesis of mesoporous materials. Reprinted with permission from [4], J. Y. Ying et al., *Angew. Chem. Int. Ed. Engl.* 38, 56 (1999). © 1999, Wiley.

Japanese scientists have proposed a lamellar transformation mechanism to explain the synthesis of mesoporous FSM-16 materials from kanemites [35]. However, this mechanism is restricted to specific conditions. From silicate oligomer precursors to the final hexagonal ordered mesoporous materials, the lamellar intermediate occurs only in a narrow pH range. Actually, it is not necessary to form the lamellar intermediate when the hexagonal mesostructure is assembled [24]. Furthermore, what is the difference in the structures between FSM-16 and MCM-41 materials has not been fully understood [36].

4. INTERACTION BETWEEN ORGANIC AND INORGANIC SPECIES

Considering the kinetic formation process of mesoporous materials, the total free energy change ΔG of such a process can be summarized by the contributions of the following interaction [26]:

$$\Delta G = \Delta G_{\text{Intra}} + \Delta G_{\text{wall}} + \Delta G_{\text{Inter}} + \Delta G_{\text{sol}}$$

where ΔG_{Intra} is due to interactions of organic and inorganic species, ΔG_{wall} is due to interactions between inorganic species, ΔG_{Inter} is due to organization of the organic surfactant molecules, and ΔG_{sol} is due to the chemical potential of the solutions during synthesis.

For the syntheses of mesostructures with a certain composition (such as silicates) in a given solvent system (usually aqueous solutions), the solution chemical potential ΔG_{sol} is similar and can be assumed to be a constant. Furthermore, at a given composition, concentration, pH, and synthesis temperature, ΔG_{wall} and ΔG_{Inter} may be varied within a small range, thus the formation of mesostructure is very dependent on the interaction of inorganic and organic species. G. D. Stucky et al. proposed a general synthesis pathway based on the charge density matching to minimize ΔG_{Intra} for a successful cooperative self-assembly process, that is, S^+I^+, S^-I^+, $S^+X^-I^+$, and $S^-X^+I^-$ (here S is the surfactant species; I is the inorganic species; and X is a counter-charged ion, such as Cl^-, Br^-, Na^+, or K^+) [24, 26]. With this concept, the synthesis of mesoporous materials was successfully carried out under acidic conditions; many mesoporous materials with new structures and new compositions have been synthesized [24, 26]. It is a real breakthrough that molecular sieves can be synthesized in acidic media.

It has been known that careful design of the headgroup of surfactant molecules is very important to the synthesis of mesoporous materials. During the formation of siliceous mesostructures under a basic condition, the inorganic precursors are negatively charged silicate oligomers (I^-); therefore, cationic surfactants such as long-chain quaternary ammonium salts should be used (S^+) to achieve a Coulombic interaction (S^+I^-) to form an ordered mesostructure. An example of S^-I^+ interaction is the synthesis of mesostructured alumina with anionic surfactants such as sodium lautyl sulfate and cationic inorganic species such as Al_{13} ($[Al_{13}O_4(OH)_{24}(H_2O)_{12}]7^+$), which has a Keggin structure similar to that of heteropolyoxometallates [37, 38]. Organic and inorganic species with the same charge can also be used to form mesostructures; however, counter-charged ions are necessary in such synthesis systems to balance the charge density, such as the $S^+X^-I^+$ ($X^- = Cl^-$, SO_4^{2-}, NO_3^-, etc.) and $S^-M^+I^-$ ($M^+ = Na^+$, K^+, etc.) synthesis pathways.

The S^0I^0 synthesis pathway was also proposed to explain the generation of mesoporous materials [39–43]. It was carried out at near-neutral conditions with the use of neutral organic amines or nonionic oligomeric surfactants (S^0) as templates. It should be noted that under such conditions both the inorganic and organic species are not strictly neutral, and hydrogen bonding may play an important role for these partially charged inorganic and/or organic species. Complexation may also occur, especially in the synthesis of nonsiliceous mesostructures; for example, niobium (V) ethoxides and long-chain alkyl amines can be used to form mesostructured niobium oxide [44]. Another combination is covalent bonding S-I interaction. For example, silicon-containing surfactants ($C_{14}H_{29}N(CH_3)_2$-$Si(OC_2H_5)_3Br$) can be used as a template in the synthesis of mesoporous silicate materials, with TEOS as a silica source (Fig. 4).

5. SYNTHESIS STRATEGY

5.1. *Hydrothermal* Synthesis

Mesoporous materials are usually synthesized by solution reactions under the *hydrothermal* conditions similar to those of zeolite molecular sieves. The synthesis undergoes a typical sol–gel process [24, 45]. Because the synthesis temperatures for mesoporous materials are relatively low (from room temperature to 100 °C), it is not the conventional hydrothermal synthesis. Mesoporous silica materials can be synthesized under basic media or strong acid conditions [24, 27]. Under alkaline conditions, due to the reversible polymerization of the silicate species [45], there is no limitation in selecting silica sources; therefore, all kinds of silica regents such as silica gels, silica sols, water glass, fused silica, and TEOS

can be used for the syntheses [12, 13]. Generally, NaOH, KOH, $NH_3 \cdot H_2O$, and organic amines can be used as the alkaline sources, and the best pH range is 11.0–11.5 [46–53]. In the reaction composites for preparing MCM-41 initially reported by the Mobil Company [12, 13], several types of mixing silica sources were employed; this might be attributed to the consideration of the amorphous inorganic walls of MCM-41. But later on, TEOS was found to be one of the most convenient and applicable silica sources [24, 27, 53]. MCM-41 with highly ordered hexagonal mesostructure is the most popular mesoporous material, which can be synthesized under a wide range of reaction conditions. High-quality mesoporous silica MCM-41 with a large pore size (>4 nm) can be obtained generally by recrystallization or mixed surfactant strategies [52].

Cubic bicontinuous MCM-48 is thought to be somewhat difficult to prepare. Since the first report of mesoporous silica materials by Mobil scientists in 1992, more than 4000 papers were published on the subject by the end of 2001. But the papers relating to MCM-48 are less than 10% of the total (<400 papers); the main reason for this is the difficulty of the synthesis of MCM-48. However, if the processing conditions are well controlled and the surfactant is properly selected (more specifically, the hydrophilic volume of the surfactant is relatively small), it is not hard to obtain cubic MCM-48 of high quality [52, 53]. It has been proved that TEOS is the best silica source for synthesizing MCM-48, and the reaction temperature is usually higher than 100 °C. Although R. Ryoo et al. reported that fused silica and other silica sources can also be used to prepare MCM-48, ethanol is still necessary [54]. Some successful examples of the synthesis of MCM-48 are as follows:

(A) Employing TEOS as a silica source; ethanol molecules produced during the hydrolysis reaction tend to invade the hydrophobic area of the surfactant micelle, leading to head volume expansion of the surfactant.

(B) Adding organic molecules with medium polarity, such as triethanolamine, to the synthesis system; these molecules are likely to stay at the outer fringe of the hydrophobic area of the micelles. In this way, MCM-48 can be obtained, even when sodium silicate is used as a silica source [55].

(C) Using special surfactants, such as Gemini surfactants, $C_{m\text{-}12\text{-}m}$ ($C_mH_{2m+1}N(CH_3)_2C_{12}H_{25}N(CH_3)_2C_mH_{2m+1}Br_2$) ($m = 12$–22) or CDBA ($C_{16}H_{33}N(CH_3)_2CH_2C_6H_5Br$); these surfactants contain the C_{12} long chain or benzene ring connected to the heads of the surfactants and have large hydrophobic volumes. Surfactant $C_{22\text{-}12\text{-}22}$ favors the formation of high-quality MCM-48, even at room temperature at low surfactant concentration (0.4 wt%) [52].

(D) Adding a small amount of negatively charged surfactants such as sodium lauryl sulfate. The ionic pairs generated by a part of positively charged and negatively charged surfactants are less hydrophilic and withdraw to the hydrophobic area of the micelles; therefore they can serve as cotemplates to form MCM-48 [56].

(E) A careful phase transformation from hexagonal MCM-41 at high temperature can also be employed to prepare cubic MCM-48 [53, 57–59].

Q. Huo et al. creatively found that high-quality mesoporous materials could be synthesized under strong acidic conditions [24, 27]. Hexagonal mesoporous material structures synthesized under acidic conditions with the use of cationic surfactant CTAB (which is also called SBA-3) are much different from MCM-41. There are many halogen ions (for example, Cl^-) in as-synthesized SBA-3 to counteract the charges of cationic surfactants. After calcination, the mesostructure (hexagonal) and chemical composition of SBA-3 are similar to those of MCM-41, but it is still unknown whether the inorganic wall structure is the same. More recently, it has been found that the walls of calcined SBA-3 have disordered micropores [60], and the reason for the generation of micropores is still unknown. Furthermore, the hydrothermal stability and the surface acidity after grafting of Al are quite different for SBA-3 and MCM-41.

Generally, TEOS is the only silica source used for the synthesis of mesoporous materials under acidic conditions, but recently it has been found that sodium metasilicate (Na_2SiO_3) can used as a substitute [61]. By fast acidification of Na_2SiO_3, low-polymerized silicate agglomerate species can be produced as well. Because of the acidic conditions, the polymerization of silicates is not reversible [45], and such oligomers are needed to act as the precursors to assembly of the mesostructure. Preparation of mesoporous materials under acidic conditions with a cationic surfactant as a template was usually carried out at room temperature, and high-quality mesoporous silica material can be produced within 3–5 min. After the precipitation of mesostructured silica, washing with water is not necessary because any extra species can be burned out in the late calcination process. With the use of large-head-group cationic surfactant, such as $C_{16}H_{33}N(C_2H_5)_3Br$ (CTEABr), as a template, a highly ordered mesoporous silica material, SBA-1, with a simple cubic structure (space group *Pm3n*) can be prepared under acidic conditions (Fig. 5) [24, 27].

A breakthrough for synthesizing mesoporous silica materials was the preparation of a new family of mesoporous silica, SBA-11, SBA-12, SBA-15, and SBA-16, with the use of nonionic alkyl poly(ethylene oxide) (PEO) oligomeric surfactants or poly(ethylene oxide)-*block*-poly(propylene oxide)-*block*- poly(ethylene oxide) (PEO-PPO-PEO) copolymers as the templates under acidic conditions [62, 63]. One of the most outstanding mesoporous silica materials is SBA-15, with its highly ordered hexagonal mesostructure (Figs. 5 and 6). The structural regularity of SBA-15 is better than that of MCM-41; the pore size is very large and tunable between 4.6 and 13 nm; the pore volume is between 0.85 and 1.3 cm^3/g. The thickness of the silica walls of SBA-15 is between 3.1 and 7.0 nm. The templates are triblock copolymers (such as P123, $EO_{20}PO_{20}EO_{70}$), have many oxygens in their composites, and therefore are easily removed [62, 63]. After calcination at 190 °C, about 99 wt% organic species can be decomposed, and after calcination at 450 °C, 100% templates can be burned off. Solvent extraction is another effective way to remove a template. In this way, the templates can be recovered and re-used to prepare SBA-15 with similar quality [62, 63]. The large-pore SBA-15 is

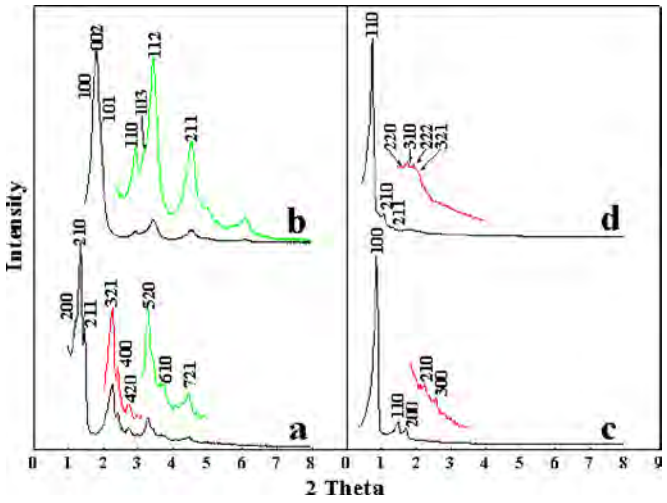


Figure 5. Representative XRD patterns of mesoporous materials. (a) SBA-1 prepared with cationic surfactant $C_{16}H_{33}N(C_2H_5)_3Br$. (b) SBA-2 prepared with Gemini surfactant. (c) SBA-15 prepared with triblock copolymer P123, $EO_{20}PO_{70}EO_{70}$, as a template at 100 °C. (d) SBA-16 prepared with triblock copolymer F127, $EO_{106}PO_{70}EO_{106}$, as a template under acidic conditions.

stable at high temperature and in boiling water. The structure of SBA-15 is still thermally stable after being calcined at 1000 °C for 24 h. Calcined SBA-15 can be stable after being heated in boiling water for more than 1 week. If a large amount of nonpolar organic compounds (for example, 1,3,5-trimethylbenzene, TMB, normally called organic-swelling agent) is added in to synthesis system, the products will be silica mesofoams [64]. The materials have a relatively narrow pore size distribution, and the pore size can be enlarged up to 40 nm; the pore volume can be as high as 2.5 cm^3/g (that is, the porosity is as high as 90%), which is comparable to that of silica aerosols. The silica mesoforms also have a high thermal stability (up to 1000 °C).

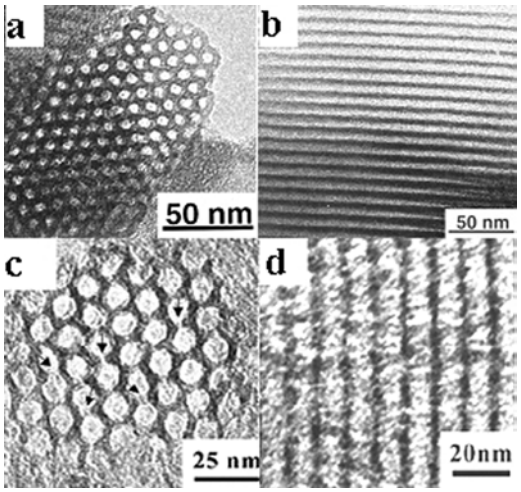


Figure 6. TEM images for (a, b) conventional mesoporous SBA-15 with a thick wall and (c, d) 3D SBA-15 with large mesotunnels on the wall. The images were recorded along the (a, c) [100] direction and along the (b, d) [110] direction.

In a similar synthesis system, different mesostructures can be obtained with different surfactants. For example, with $C_{16}H_{33}EO_{10}$ (Brij56), cubic SBA-11 (space group *Pm3m*) can be formed, whereas with $C_{18}H_{37}EO_{10}$ (Brij76), SBA-12 is obtained [63]. With the use of triblock copolymer with a high EO/PO ratio (such as triblock copolymer F127, $EO_{106}PO_{70}EO_{106}$, F98, $EO_{123}PO_{47}EO_{123}$, F108, $EO_{132}PO_{50}EO_{132}$) [65], highly ordered, large-pore cubic caged SBA-16 (space group *Im3m*) is produced (Fig. 5). Another example, FDU-1, is prepared with poly(ethylene oxide)-*block*-poly(buthylene oxide)-*block*-poly(ethylene oxide) (PEO-PBO-PEO) copolymer, B50-6600 ($EO_{39}BO_{47}EO_{39}$), which contains hydrophobic segment BO, $[CH_3C(CH_3CH_2)O]_n$ blocks [66]. With a block copolymer with a low EO/PO ratio such as L121, $EO_5PO_{70}EO_5$, unstable lamellar mesostructured materials are generally obtained. It has been suggested that the formation of the mesophase templated by block copolymers is processed through the $(S^oH^+)(X^-I^+)$ approach. With the addition of inorganic salts such as NaCl etc. to introduce Cl^- into the synthesis system, it takes less time for silica to precipitate and form a mesostructure [63, 67, 68].

Highly ordered mesoporous materials (SBA-15) have attracted more and more attention not only because they have high-quality structural regularity, thick inorganic walls, and excellent thermal and hydorthermal stability, but also because the template is economically cheap and nontoxic, and the synthesis is quite simple and reproducible [62, 63]. The pore size of the materials can be tuned easily through hydrothermal treatment; that is, a higher heating temperature achieves a larger pore size [62, 63]. More importantly, it has recently been revealed that the wall structure of SBA-15 is quite different from that of MCM-41. A large number of disordered micropores are distributed within the wall of SBA-15; even mesopores with diameters between 2 and 3 nm can be observed [69–71]. The detailed structure of these randomly distributed micropores is still unknown, but it is believed that these micropores may result from the partial occlusion of PEO segments into the inorganic walls during the synthesis process. In fact, this was already realized during the first synthesis of SBA-15 [62, 63]. When hydrothermally treated at high temperature, PEO segments of the triblock copolymers become more hydrophobic and draw back from the inorganic walls, which makes the core of the micelles of the triblock copolymers larger and results in the formation of SBA-15 with a larger pore size. The generation of the small number of mesopores on the silica walls of SBA-15 may not simply be attributed to the insertion of PEO groups in the inorganic walls; the transmutation of micelles during the hydrothermal treatment may be another important factor. The mesopores do not shrink because of the withdrawal of PEO segments from the walls at high temperatures; instead, such mesopores become larger. Zhao and co-workers have demonstrated a new method for the preparation of 3D SBA-15 (Fig. 6) [72]. TMB was added into the synthesis batch as soon as the SBA-15 embryo generated, followed by a hydrothermal treatment at high temperature (100–150 °C). With such a process, the core of the micelles, comprising the block copolymers and TMB, is swollen and breaks through the original micropore area of the silica walls, and finally large-sized disordered mesopores

(with diameters as large as 8 nm) are produced (Fig. 6). Almost no micropores were observed on the inorganic walls of these 3D SBA-15 materials, and all of the pores in the walls were large mesopores. N_2 adsorption/desorption isotherms show that because of the interconnection of the ordered channels by the mesopores on the walls, the calculated pore diameter can be as large as 22.8 nm, and TEM images confirm that the pore structure is like a 3D connected honeycomb (Fig. 6).

The existence of micropores on the silica wall of SBA-15 may favor their applications in catalysis and separation [73, 74]. The synthesis of ordered mesoporous carbon is especially advantageous because of this structural property [75–82]. But for some applications, these micropores are not required, and SBA-15 without micropores can be prepared via a nonaqueous system [65]. Conventional SBA-15 materials were also used as hard templates to synthesize semiconductor nanowires, and isolated nanowires can rarely be obtained [83–85].

5.2. Nonaqueous Synthesis

Nonaqueous synthesis is a convenient method for the preparation of ordered mesoporous materials, and this method has become more and more acceptable and profitable recently [86–90]. The nonaqueous strategy was first used by Brinker, an American sol–gel chemist, in the preparation of mesoporous silica films [91]. Zhao and co-workers further developed this method to synthesize large-dimension mesoporous silica materials [92, 93] and nonsiliceous mesoporous material [86, 90]. This nonaqueous synthesis method adopted in typical sol–gel chemistry proved to be successful, especially when the surfactant self-assembly was introduced into this system [45]. TEOS dissolved in the organic solvent (normally ethanol, tetrahydrofuran, acetonitrile) was prehydrolyzed with a stoichiometic quantity of water at a temperature of 25–70 °C with a small amount of acid (usually HCl) as a catalyst. Then low-polymerized silicate species that were generated by hydrolysis could interact with the surfactant. During the solvent evaporation, the silicate species further polymerized and condensed, forming the mesostructure due to the fast self-assembly of silicates and surfactants. Finally mesoporous materials can be obtained by calcination to remove the surfactants. Nonaqueous synthesis is a versatile method for preparing mesoporous films [91–97], monoliths [98–101], and spheres [102]. Compared with the "hydrothermal" synthesis method, nonaqueous synthesis is more convenient and faster. From the XRD patterns, the diffraction peaks at a relatively high angle range ($2\theta = 3$–$5°$) can hardly be observed for the mesoporous materials prepared by the nonaqueous method, and it seems that the regularity of mesostructures prepared by this method is poorer than that from the "hydrothermal" synthesis method. However, careful observations from TEM show that the structural regularity of the mesoporous materials prepared through the nonaqueous synthesis method is pretty good, and nearly all domains are highly ordered [65]. There may be fewer and broader diffraction peaks because the mesostructure forms too fast in the nonaqueous system, which may lead to nonuniformity of the silica gel particles. Furthermore, the surface area of the materials prepared by the nonaqueous route is lower than that prepared by "hydrothermal" synthesis. Although the recrystallization process can be used to improve the quality and increase their surface areas, why the surface area is so low is still unknown. In addition, another difference between the materials derived by nonaqueous synthesis and "hydrothermal" synthesis is that there are no or fewer micropores in the silica walls of the mesoporous silica products templated by the triblock copolymer, such as P123 under nonaqueous conditions. It is difficult to tune the thickness of the inorganic wall by the nonaqueous synthesis route compared with the preparation of SBA-15 by the hydrothermal method, and the maximum pore diameter that the block coploymers produced is achieved in nearly one step [92].

The nonaqueous synthesis strategy has proved to be one of the simplest and easiest methods for preparing 3D cubic mesoporous silica materials [65, 88, 89, 92]. With the use of block copolymers with long PEO segments, such as F127, F108, F98, or surfactant blends, mesoporous silica materials with a structure similar to that of SBA-16 (*Im3m*) can easily be synthesized [92]. Adding a small amount of organosilicates, for example, mercaptopropyltrimethoxysilane (MPTS), or a small amount of nonpolar organic molecules, for example, ethylbenzene, toluene, TMB, etc., at room temperature under acidic conditions, Zhao and coworkers have synthesized a bicontinuous cubic mesoporous silica material, FDU-5 [89], that has an *Ia3d* structure identical to that of MCM-48. The pore size (up to 10 nm) of the material is larger than that for that of MCM-48, and the pores of FDU-5 directly connect with organic function groups. These large-pore mesoporous silica materials (FDU-5) have been used as hard templates to prepare mesoporous carbon materials with a bicontinuous cubic structure (space group *Ia3d*) [89].

The nonaqueous synthesis route was further extended to prepare nonsiliceous mesoporous materials by Yang et al. [86, 90]. The synthesis process is quite simple. In general, metal halides, such as $NbCl_5$, $TiCl_4$, $AlCl_3$, $SiCl_4$, and so on, were dissolved in ethanol, and then, after the addition of block copolymer surfactants, mesostructures could be obtained by completely evaporating the solvent at 40–60 °C. Mesoporous materials can be obtained via removal of the surfactants by calcination at 300–450 °C. The materials prepared by this method were thermally stable (generally stable at ~500 °C), and the pore size distribution was relatively uniform, but the surface area was relatively low (<200 m^2/g) [86, 90]. It is proposed that this synthesis method is a sol–gel process with an acid-controlled mechanism. MCl_x can react with ethanol by alcoholysis reaction and produce intermediate $M(OEt)_nCl_m$ species that have a low polymerization speed; then mesostructure is formed by slow hydrolysis with the moisture in air [86, 90]. The coordination bond between the oxygen atoms of block copolymers and the metal ion is believed to be important to the formation of final organic/inorganic composite mesostructures.

Since the interaction between inorganic species is less considered, mesoporous oxide materials such as ZrO_2 and Nb_2O_5 and some metal oxides containing mesoporous silica are ordered, and some others are disordered mesostructures [86, 90]. Recently Zhao and co-workers demonstrated a new

"acid-base pair" concept, in which the interaction between inorganic species is considered and the inorganic sources can be classified into acids or bases by fundamental chemical principles. With the selection of proper inorganic precursors by this concept, a large family of mesoporous metal oxides, metal phosphates, mixed metal oxides, mixed metal phosphates, and metal borates can be successfully prepared. For example, highly ordered hexagonal (*p6mm*) and cubic mesoporous $TiPO_4$ can be obtained (Fig. 7) by this route. The materials obtained have a uniform composition distribution and are thermally stable (>700 °C in general) and highly ordered, and the pore size distribution is very narrow (FWHM < 3 nm).

5.3. Synthesis Rules

In general, the phase of mesoporous materials is very dependent on some factors such as reaction temperature, the concentration and the packing parameter of the surfactant molecules, and so on. It was assumed that the mesostructure is determined only by the concentration of the surfactant, and this observation is based on the fact that the hexagonal liquid crystal phase is first formed at low surfactant concentration, whereas the transformation from the hexagonal to the cubic mesophase occurs at higher concentrations. Later, the packing parameter (g) was proposed as a guideline to predict and explain the formation of the mesoporous product phase [29].

In $g = V/a_0 l$, V represents the hydrophobic volume of the whole surfactant molecule; a_0 represents the effective area of the surfactant head group, and l is the length of the surfactant chain. Although it is the result of a simple geometric computation, the value of g can well describe the type of liquid crystal phase formed under specific conditions. As for the synthesis of mesoporous materials, it can enlighten us about how to control the reaction conditions and parameters to obtain the expected mesophase. Moreover, it provides an excellent explanation of the experimental phenomena observed. When g is smaller than 1/3, caged SBA-1 (*Pm3n*, cubic) and SBA-2 ($P6_3/mmc$, 3D hexagonal) form. When g is between 1/3 and 1/2, MCM-41 (*p6mm*, 2D hexagonal) forms. When g is between 1/2 and 2/3, MCM-48 (*Ia3d*, cubic) forms. When g is near 1, a lamellar mesostructure can be formed.

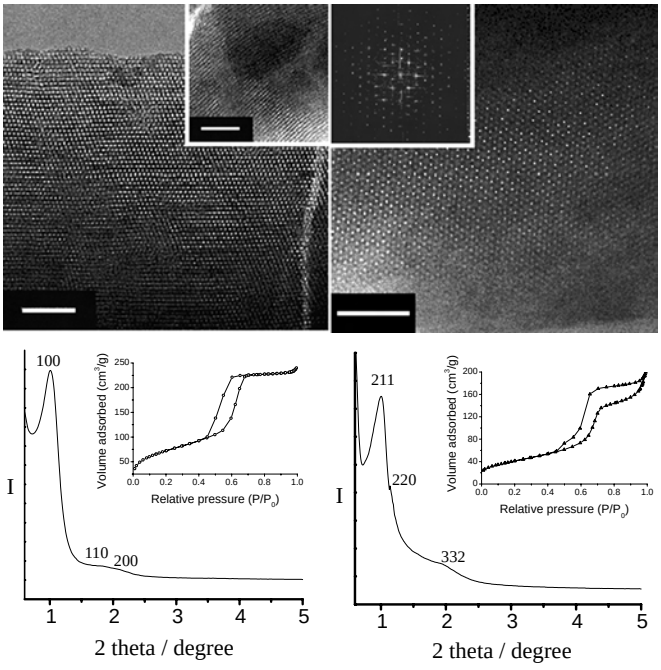


Figure 7. Structural characterizations of mesoporous titanium phosphate (TiPO). (a) TEM images (viewed from [110] and [100] directions). (b) TEM image, Fourier diffractogram (both from the [111] direction). (c) XRD, N_2 sorption curves of 2D hexagonal mesoporous TiPO. (d) XRD, N_2 sorption curves of bicontinuous 3D cubic mesoporous TiPO.

Under similar synthesis conditions, the influence of the reaction conditions and the inorganic species is nearly the same for the cooperative assembly of the surfactant-inorganic species. In this case, the properties of the surfactant (geometry, charge, and structure) have a large effect on the assembly of the inorganic-organic mesostructure; therefore different mesophases can be obtained with the use of different surfactants. In other words, the mesostructure of the products can be controlled by the selection of proper surfactants or the addition of auxiliary agents that have an influence on the packing parameter of the surfactants [24, 27, 52]. Some examples are shown as follows:

(A) Cubic SBA-1 can be prepared directly by alkyltriethylamine ($C_nH_{2n+1}N(C_2H_5)_3Br$, n = 10–18) [52], because the g value is decreased by the large head group of the surfactants. The lamellar mesostructure can be directed by bi-tail surfactants (e.g., $(C_{12}H_{25})_2N(CH_3)_2Br$) because they have a large hydrophobic part. The g value increases with the V value.

(B) The length of the hydrocarbon chains of the surfactants has a little influence on the mesostructure formed. The reason for this is that the value of V/l undergoes almost no changes with different lengths of hydrocarbon chains; that is, the g value does not change. But when the hydrocarbon chains are long enough (e.g., longer than 20 carbon atoms), the carbon chains can be wound, and thus the V value increases a little while l decreases, and the g value increases. Therefore, the lamellar mesostructure can easily be templated by surfactants $C_nH_{2n+1}N(CH_3)_3Br$ (n = 20, 22).

(C) The Gemini surfactant $C_{m\text{-}s\text{-}m}$ is a kind of special surfactant and consists of two head groups connected by a hydrocarbon chain. Controlling the length of the hydrocarbon chains can vary the area of the surfactant head groups. For example, with the surfactant $C_{16\text{-}s\text{-}16}$, $C_{16}H_{33}N(CH_3)_2(CH_2)_sN(CH_3)_2C_{16}H_{33}Br_2$ as a template under base conditions, when s is changed from 2 to 12, the product obtained was transformed from the lamellar to the hexagonal phase (MCM-41), and the final products (when s = 12) were cubic mesophase (MCM-48).

(D) Many experimental phenomena, including the influence of the organic additives on the synthesis, the mesophase transformation, and so on, can be explained based on the g value [103].

Although the g value is a wonderful parameter for the evaluation of the mesoporous product phase and explains well the experimental phenomena observed from cationic surfactants, somehow it does not apply very well to block copolymer synthesis systems. Recently, we found that the critical micelle concentration (CMC) of block copolymer surfactants is a powerful criterion for the experimental design of the synthesis of ordered mesoporous materials. When the CMC of the block copolymer is below 20 mg/liter at room temperature, the template is ready to synthesize ordered mesostructures. When the CMC is approximately between 20 and 300 mg/liter, some strategies can be used to decrease the CMC to give ordered mesostructures, such as increasing the temperature and/or ionic strength of the reaction liquors. When the CMC is further increased (>300 mg/liter), it will be difficult to make ordered mesostructures.

5.4. Secondary Synthesis

Besides the above-mentioned direct synthesis methods, secondary synthesis is also one of the most important synthesis strategies [104, 105]. To increase the wall thickness or local structure regularity, secondary synthesis, that is, grafting or recrystallization, must be employed. For example, MCM-41 can be stabilized in boiling water by the reaction of MCM-41 and $AlCl_3$ [106]. Compared with the matrix, the materials after the treatment have better mechanical and hydrothermal stability. This may be attributed to the improvement of wall thickness and internal silica linkage, and mend of some structural defects. After water vapor treatment at high temperature, the materials still have a strong Bronsted acidity on the surface.

Recrystallization can make the structural regularity of mesoporous materials more perfect and ordered; however, only a few research groups have comprehended this trick. Most researchers have employed hydrothermal treatment to improve the structural regularity [52]; to be specific, the solid mesoporous materials obtained at room temperature were put into an autoclave within the mother solution, where they underwent hydrothermal treatment at high temperature (100–150 °C). Differing from the above treatment, the recrystallization process includes the addition of water to an as-synthesized mesoporous material, which is then heated at 100 °C in neutral medium for several days (about 1 week). The quality (structural regularity, thermal stability, etc.) of the materials can obviously be improved after this treatment; sometimes the pore size can also be increased. Moreover, secondary synthesis is also a good method for preparing mesoporous material-doped transition metals [49, 50].

Another secondary synthesis method has been demonstrated by Zhao and co-workers to synthesize a highly ordered hexagonal mesoporous silica structure with the use of a nonionic oligomeric surfactant with long hydrophilic chains [104]. This process involved the adsorption of an extra silica source (TEOS) into the directly synthesized products and secondary hydrolysis process in either base or acidic media. The treatment resulted in great improvement of the ordering and hydrothermal stability of the mesoporous silica materials. Similar secondary synthesis procedures have recently been applied in the synthesis of high-quality mesoporous films [105].

6. THE DIFFERENCES WITH TRADITIONAL ZEOLITES

There are many differences between mesoporous materials and traditional zeolite molecular sieves. As far as the structure is concerned, the differences are:

(A) Different crystallinity. Mesoporous materials are imperfect crystals, and at the atomic level they cannot be defined as crystals.

(B) Structure codes. Each of the zeolite structures is given a structure code by the International Zeolite Association (IZA). For example, the code of zeolite faujasite Y is FAU [107]. But for mesoporous materials, it is difficult to give an IZA structure code, because there are so many differences in properties between the mesoporous materials with identical structure and composition.

(C) Building blocks. The traditional zeolite molecular sieves are mainly constructed by tetrahedrons (TO_4), but the inorganic walls of the mesoporous materials are amorphous, and many coordination types, such as TO_6 and TO_5, are allowed as the building blocks to construct a mesoporous material structure.

(D) From the viewpoint of the atomic level, the connections between the inorganic species of the frameworks for mesoporous materials are relatively weak. For example, the four corner points of the silicon-oxygen tetrahedron are not necessarily connected with other tetrahedrons.

(E) In general, mesoporous materials (especially silica) are not stable in water (referred to as "hydrothermally" unstable). After being heated for 6 h in boiling water (100 °C), calcinated MCM-41 becomes amorphous and loses its structural regularity [63, 108]. Al-containing MCM-41 is more stable than pure silica MCM-41 in hot water [106]. Mesoporous silica materials synthesized with nonionic oligomeric surfactants or block copolymers (SBA-11, -12, -15, -16) are far more stable and can retain their structure after being heated in boiling water at least for 10 days. People have recently tried to increase the framework stability by various methods, including the use of zeolite precursors as the inorganic sources [109–112]. It should be noted that stability in boiling water is not true hydrothermal stability. The true hydrothermal stability should be studied under the treatment with 100% water vapor at 600–800 °C. Until now, no detailed paper on the hydrothermal properties of mesoporous materials has been reported [112]. But based on the current literature, the structure of mesoporous silicates (or aluminosilicates) is destroyed after 100% water vapor treatment for 4 h at 800 °C [112].

As far as the synthesis is concerned, there are also many differences between mesoporous materials and traditional zeolite molecular sieves. First, the synthesis temperature is much lower for mesoporous materials. In general the temperatures vary from room temperature to 80 °C, and mesoporous materials can even be prepared at −20 °C [113, 114]. From this point of view, the synthesis method employed

for mesoporous materials is not the classical hydrothermal synthesis technology. The hydrothermal treatment temperature for mesoporous materials is normally lower than 160 °C (commonly 80–150 °C) even after these materials precipitate (this means the mesoporous materials have been formed), otherwise, crystalline microporous material zeolite can be formed at high temperature (>150 °C) [113]. Second, the formation rate of mesoporous materials is very fast. Mesostructure forms in several minutes, even in several seconds. In contrast, it generally takes several hours or days to prepare microporous zeolite materials. Third, the inorganic walls of mesoporous materials are amorphous; therefore their composition can be varied, and the inorganic framework does not have to be composed by tetrahedron building blocks. In this context, mesoporous materials can be made up of any inorganic elements. Because of this consideration, large amounts of nonsiliceous mesoporous materials, metal oxides, metal sulfides, metal phosphates, and even metal mesostructure have been prepared. Fourth, mesoporous materials can be synthesized under true nonaqueous conditions, such as preparing mesostructure through rapid solvent evaporation, with alcohol as a solvent. However, molecular zeolite sieves are still synthesized in pseudononaqueous conditions [115]. A little water is necessary for the formation of zeolite structure. Sometimes the structure of zeolite crystals may be destroyed by completely evaporating the solvent in an autoclave. Fifth, the morphology of mesoporous materials is relatively easy to control. In some cases, the morphology obtained is not related to the mesoporous material structure. Mesoporous materials can be prepared with a morphology of spheres [97, 116–120], fibers, and so on.

7. COMPOSITION

Metal doped mesoporous silica materials have attracted considerable attention because they show potential applications as hosts, adsorbents, and catalysts [73, 74]. Substituted MCM-41 and SBA-n mesoporous materials can be prepared by direct hydrothermal synthesis. The content of doped metals for substituted MCM-41 and MCM-48 prepared by direct synthesis is usually below 10 wt%. Mesoporous TiO_2 [88, 121], ZrO_2 [122], Al_2O_3 [111, 123], Ga_2O_3 [38], $AlPO_4$ [49, 124], $TiPO_4$ [125], and some other nonsiliceous materials have been synthesized as well. In addition to the previously described nonsiliceous materials prepared via the nonaqueous synthesis route, many cases have been reported for the preparation of nonsiliceous materials in aqueous solution, normally with long-chained organic amines or cationic surfactants as the templates. Most of these products are unstable; additional methods should be employed to stabilize the resultant mesostructures.

Adding a large amount of doped metal ions into a synthesis batch for mesoporous silica materials can result in a severe decrease in the mesostructural regularity of the materials. In fact, a large amount of transition metal ions can be introduced into mesoporous materials by a secondary synthesis that is similar to the impregnation method for preparing catalysts. Because the inorganic walls of mesoporous materials are amorphous, it is meaningless to try to determine whether the substituted atoms are in the inorganic framework of mesoporous materials. There is almost no doping quantity limitation for secondary synthesis. Based on currently published papers, the differences in some properties between the doped metal mesoporous materials prepared by the direct method and those prepared by the secondary method are not clear. For example, Al-MCM-41 prepared by secondary synthesis shows a better stability and surface acidity than that from direct synthesis. Heteropolyoxometallates, $H_2W_{12}PO_4$, can be incorporated into a mesoporous material (SBA-15) to prepare a good solid acid catalyst. When the doped amount is higher than 150 wt%, the crystalline phase of $H_2W_{12}PO_4$ cannot be detected by X-ray diffraction (XRD).

Mesoporous carbons are a kind of new nonsiliceous material discovered recently by Korean scientist R. Ryoo and co-workers [75–82]. The materials have an ultrahigh surface area (as high as 2500 m^2/g) and a very large pore volume (up to 2.25 cm^3/g) [126], which makes them promising materials for catalyst supports, hydrogen storage, and electrodes, and therefore they attract much attention [127, 128]. With ordered mesoporous silica materials as the templates, amorphous carbons can be incorporated into their channels, and ordered mesoporous carbon can be obtained after removal of the silica framework by HF or NaOH. The carbon materials have a structure that is a replica of their silica templates, which means the two materials belong to the same space group, but the pores and frameworks are inversed from that of the mesoporous silica. A series of mesoporous carbons have been prepared and named CMK-n (carbon materials of Korea). Mesoporous carbon derived by SBA-15 has a 2D hexagonal mesostructure (*p6mm*) and is called CMK-3. Because of the micro- and mesopores in the inorganic wall of SBA-15, after removal of silica, amorphous carbons filled in the micropores act as pillars to support the carbon rods formed in the 1D channel of SBA-15; therefore a hexagonal mesostructure is constructed. Mesoporous carbon templated by bicontinuous cubic MCM-48 is called CMK-1 or CMK-5. Because of the 3D interconnected pore structure of MCM-48, the CMK-5 obtained (CMK-1) is composed of continuous amorphous carbon frameworks, and the space group is a little different from that for MCM-48. Ordered mesoporous carbon cannot be templated by MCM-41, and only amorphous carbon nanorods or nanowires can be obtained. Because MCM-41 prepared under base condition has a solid inorganic wall without micropores, the pillars that support the carbon rods cannot be formed like that in the case for SBA-15, and after the removal of silica, the carbon rods or wires will be disassembled. The pore size control of mesoporous carbons is relatively difficult, and mesoporous silica with large pores may not produce large-pore mesoporous carbon. The only way to accomplish this is to control the wall thickness through different kinds of mesoporous materials. Recently, Ryoo and co-workers employed SBA-3 as a template, which has a few micropores on its inorganic wall, and successfully prepared mesoporous carbon with a small pore size [60]. Y zeolite was also used as a template to synthesize nanoporous carbon molecular sieves by a similar method [83].

There are several alternative methods for incorporating carbon into the channel of mesoporous silica materials,

mainly including chemical vapor deposition (CVD) [82] and the solution method [75–82]. The former is relatively difficult to control, while the latter is more effective. In a typical synthesis, sucrose or furfural alcohol is employed as a carbon source by impregnation. Carbonization was carried out at 90 °C with concentrated H_2SO_4. The mesopores of the template can be completely filled with the carbons by the process described above, repeated one or two times. After removal of silica by HF or NaOH, mesoporous carbon is obtained. Attention has to be paid to this step to ensure that the silica is completely removed.

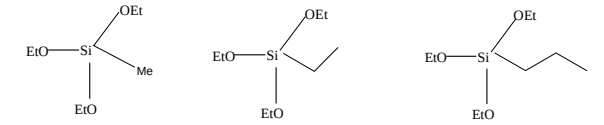
At present, the synthesis of spherically structured mesoporous carbon has encountered difficulty. With mesoporous silica with a cage structure as a template, the windows connecting the cages are too small (the entrance size is smaller than 4 nm), so the replicas of carbon spheres are not well connected and easy to take apart; therefore the resulting mesoporous carbon is poorly ordered. Zhao and co-workers found recently that 3D spherically structured carbons can be synthesized by caged cubic mesoporous silica materials with large entrance sizes, which are synthesized by using high-temperature hydrothermal treatment in the presence of inorganic salt and nonpolar organic molecules (such as TMB).

Ryoo and co-workers carefully coated the walls of the channels of SBA-15 with a layer of amorphous carbons through a well-controlled process and obtained ordered carbon tubes after the removal of silica. This new interesting carbon material is thermally stable because the carbon rods, like pillars, support the tube (Fig. 8). What's more interesting is that the material has bimodel pore size distribution, namely two sets of mesopores with uniform pore size distribution.

Mesoporous carbons are very stable, and their structural regularity is retained even after calcinations at 1600 °C (in an inert atmosphere). Still, mesoporous carbon is difficult to graphitize. Although the partly ordered lamellar structure of graphitized carbon can be observed in some small domains, a material with a complete graphitized structure has not been obtained yet, and scientists are endeavoring to achieve this target.

8. FUNCTIONALIZATION

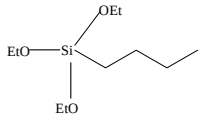
Mesoporous silica materials can easily be functionalized since their inorganic walls have abundant hydroxy groups that easily react with active organic groups such as R-Si(OEt)$_3$, R-SiX$_3$ (X = halide Cl, Br, etc.) [129–133]. The organic compounds commonly used for the functionalization are listed below. The functionalization can be attained by a simple process. In general, calcined mesoporous silica was heated in a vacuum at 100–130 °C to remove adsorbed water and then reacted with active organic compounds in an inert solvent (benzene, toluene, or cyclohexane). After filtration, the functionalized mesoporous silica product was obtained. The functionalization not only can change the hydrophilic/hydrophobic property of mesoporous materials (functionalized organic compounds are normally hydrophobic), but also takes the function of organic groups in catalysis, luminescence, separation, and sorption of heavy metal ions in environmental application. J. Liu and co-workers found that mesoporous silica materials functionalized with a mercapto group have excellent sorption ability for heavy metal ions and can be applied in wastewater treatment [134–136].



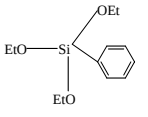
methyltriethoxysilane

ethyltriethoxysilane

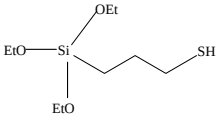
propyltriethoxysilane



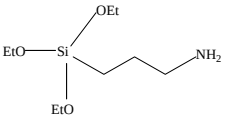
butyltriethoxysilane



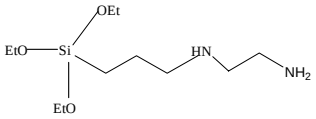
phenyltriethoxysilane



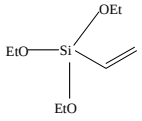
(3-mecapto)propyltriethoxysilane



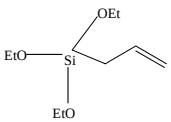
(3-amino)propyltriethoxysilane



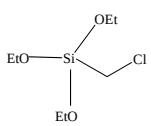
N-(2-aminoethyl)-3-aminopropyltriethoxysilane



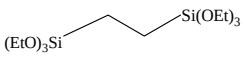
vinyltriethoxysilane



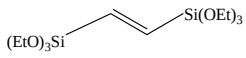
allyltriethoxysilane



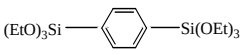
chloromethyltriethoxysilane



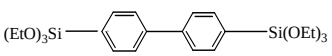
1,2-bis(triethoxysilyl)ethane



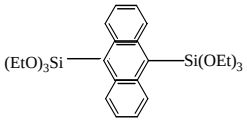
1,2-bis(triethoxysilyl)ethene



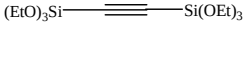
1,4-bis(triethoxysilyl)benzene



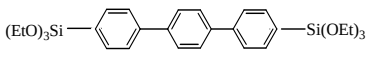
1,4-bis(triethoxysilyl)biphenyl



9,10-bis(triethoxysilyl)anthracene



1,2-bis(triethoxysilyl)ethyne



4,4'-bis(triethoxysilyl)terphenyl

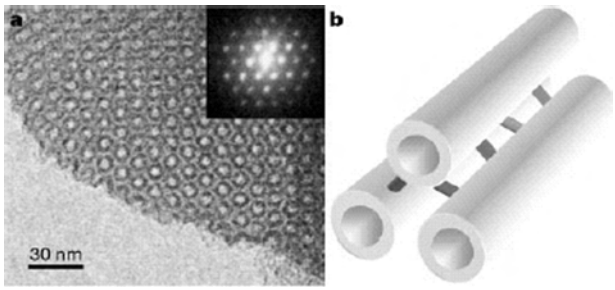


Figure 8. TEM image (a) and schematic model (b) of mesoporous carbon tubes synthesized with SBA-15 silica as a template and furfuryl alcohol as a carbon source. Reprinted with permission from [77], S. H. Joo et al., *Nature* 412, 169 (2001). © 2001, Macmillan Magazines Ltd.

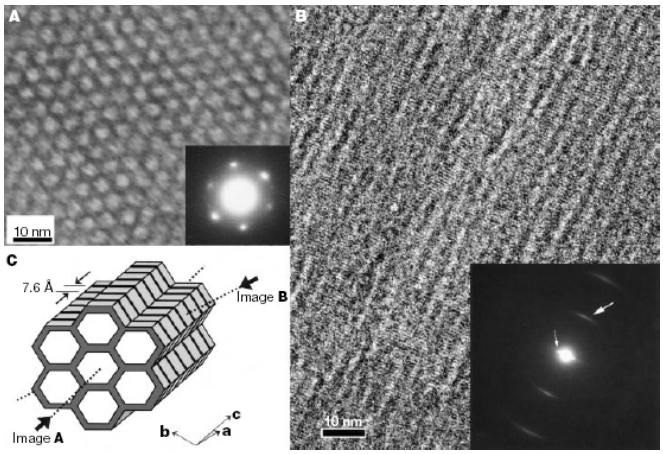


Figure 9. TEM images and electron diffraction patterns taken with (A) [001] incidence and (B) [100] incidence. (C) Schematic model of mesoporous benzene ± silica. Reprinted with permission from [138], S. Inagaki et al., *Nature* 416, 304 (2002). © 2002, Macmillan Magazines Ltd.

Another functionalized process is the direct synthesis method, in which a silica source with organic functional groups, such as $(EtO)_3Si\text{-}R$, is added directly to the silica sols of the synthesis batch [133, 134]. Because the R group cannot connect with the other tetrahedral silicate species, the quantity of the organic functional group introduced into the materials is limited. The structural regularity of the functionalized mesoporous material products is good enough when the $(EtO)_3Si\text{-}R/TEOS$ ratio is less than 10% (the largest ratio should be lower that 20%). There is still no clear determination of the differences, if any, between the functionalized materials obtained with the direct method and secondary synthesis.

Recently, Inagaki and co-workers synthesized a new kind of organic-inorganic periodic mesoporous organosilica (PMO) hybrid material with R-bridged (R represents organic function groups) organosilane $(EtO)_3Si\text{-}R\text{-}Si(OEt)_3$ as a silica source [137–140]. Different from the above functionalized mesoporous materials, the organic groups are located in the middle of the inorganic walls. Both ends of the silicate precursor molecules are able to hydrolyze and polymerize and do not affect the formation of the inorganic framework. One hundred percent of $(EtO)_3Si\text{-}R\text{-}Si(OEt)_3$ can be used as a silica source in the synthesis of mesoporous materials [141, 142]. However, it should be pointed out that the length of the hydrocarbon chain cannot be too long (shorter than C_6), otherwise ordered PMO would not be obtained. The organic function group in PMO is thermally stable (normally to 400 °C), so it is easy to remove the surfactants by either calcinaion or extraction. The PMO obtained has a high surface area (>800 m^2/g) and a large pore volume (>0.7 cm^3/g). What is more noticeable is that when benzene-bridged organosilane was used as a silica source, the resulting PMO has a crystal-like wall structure and the benzene groups can stack in a layered structure (~6 Å) by π-π interaction (Fig. 9) [138]. These ordered organic-inorganic hybrids further assemble to form an ordered mesostructure, which turns into a kind of hierarchically ordered PMO with considerable potential applications.

9. PORE SIZE

The methods used to tune the pore size of mesoporous materials are mainly as follows:

(1) Hydrothermal treatment: the higher the temperature, the larger the pore size [24, 63].

(2) using alternative surfactants [52, 63]. In general, the pore size of mesoporous materials is determined by the volume fraction of the hydrophobic part of the surfactant/inorganic composite micelles [63]. As for cationic surfactants, the longer hydrocarbon chain yields a larger pore size [12, 13]. However, it is not unlimited; when the hydrocarbon chain is longer than C_{22}, the surfactants are insoluble in water and cannot be used as templates for the synthesis of mesoporous materials. As for block copolymers, the pore size of mesoporous materials is determined by the molecular weight of the hydrophobic block. It seems that a larger molecular weight yields a larger pore size, and the pore size is not related to the total molecular weight of the block copolymers. Sometimes, using mixed surfactants as the templates can increase the pore size.

(3) Adding an organic swelling agent [12, 63]. This is an effective method of increasing mesopore size. In the synthesis of mesoporous materials with block copolymers under acidic conditions, the pore size can be as large as 40 nm when TMB is added as an organic swelling agent. Mobil Company scientists have claimed [12, 13] that the mesopore size of MCM-41 can reach 10 nm when CTAB is used as a template and TMB is added. It is noted that generally the structural regularity is decreased by the addition of TMB. The pore size of ordered SBA-15 can be increased to 20 nm [72]; for MCM-41 materials the upper limit is 6 nm, otherwise the ordered structure will be mostly destroyed. Nonpolar organic molecules, such as toluene, benzene, xylene, cyclohexane, and long-chain hydrocarbons, can be used as

organic additives to increase the pore size. Based on current reports, TMB is one of the most effective organic additives.

Sorted by the pore size of the products, the methods for achieving pore sizes are as follows:

2–5 nm: utilizing surfactants (including long-chain cationic quaternary ammonium surfactants and neutral organic amines) with various chain lengths as templates [39].
4–7 nm: employing long-chain quaternary ammonium surfactants as the templates and high-temperature hydrothermal treatment [52, 143].
5–8 nm: using cationic quaternary ammonium surfactants and adding organic swelling agents (TMB, middle, or long-chain amines) [144].
2–8 nm: utilizing nonionic oligomeric surfactants as the template [63].
4–20 nm: utilizing block copolymers as the template [62, 63].
4–11 nm: secondary synthesis (water or amines postsynthesis treatment) [144].
10–40 nm: utilizing block copolymers as templates and adding organic swelling agents (TMB) [62–64].

10. CHARACTERIZATION

Because of the potential applications of the mesoporous materials in advanced technologies such as catalysis, separation, microreactors, electronics, electrodes, microdevices, sensors, and so on, all kinds of modern characterization techniques, including XRD, TEM [82, 83, 145, 146], physical adsorption [69, 70, 81], ultraviolet diffuse reflectance spectroscopy (UV-DRS), solid-state nuclear magnetic resonance (solid-state NMR) [99], electron spin resonance (ESR) [57, 147, 148], X-ray photoelectron spectroscopy (XPS), etc. have been used to characterize the structure and formation mechanism of mesoporous materials. The most effective and essential methods are XRD, TEM, N_2 adsorption isotherm, and solid-state NMR.

10.1. XRD

It is very hard to obtain large single crystals of mesoporous materials (currently, the size of single-crystal mesoporous silica is between 1 and 2 μm); therefore, powder XRD is one of the most effective tools for characterizing mesoporous materials. Because mesoporous material structures have long-range periodicity and a large cell parameter, the diffraction peaks are usually observed at low angles (2θ is normally in the range of 0.8–8°). Measurement of low-angle XRD patterns must be carried out with caution. It is suggested that narrow slit setting, low X-ray power, and slow scan speed are the recommended measurement conditions; otherwise, artificial diffraction (or reflection) peaks will be observed, which often cause the wrong assignment of the periodicity of mesoporous materials. The limit of low-angle XRD that can be measured may be as small as $2\theta = 0.4$–0.5, depending on the type of XRD machine and measurement conditions; however, the absolute value obtained at such a low angle, even it is a true diffraction peak, may cause a relatively large error because no standard materials can be used to adjust the XRD machine at small angles. If necessary, a small-angle X-ray scattering technique should be employed to collect more accurate data at low angles. It is relatively easy to confirm a known mesophase from XRD patterns. Typical XRD patterns of some mesostructures are shown in Figures 2 and 5. Through simple comparison of XRD patterns and d values, generally, a known mesophase can be confirmed. It should be noticed that the errors for d values in a small angle range are relatively large; therefore, in addition to XRD patterns, the ratios of those precisely measured d values should also be in accordance with the known mesophases. To solve the structure of a new mesophase from the XRD pattern is a tough job. In addition to the errors of small-angle XRD, there are normally less than three diffraction peaks in the XRD patterns of the mesoporous materials. Therefore it is difficult to determine the space group from the indexation of XRD diffraction peaks of the mesoporous materials. However, it is possible as well to get information on space groups by the combination of high-quality XRD diffraction peaks with the liquid crystal phase of the surfactant. It is necessary to understand the new mesostructure of a material utilizing XRD combined with HRTEM images [145, 146]. Furthermore, a coarse judgment of the quality and structure regularity can be made by using the XRD peak width and relative intensity. Similarly, it should be noticed that XRD peaks are caused only by periodic crystal structures and represent the aggregate of the whole sample. Amorphous materials have no diffraction peaks and do not influence the XRD peaks. In other words, XRD measurements could not provide us with any information about the amorphous matters. Comparison of the periodicity of the materials by the relative intensities of the diffraction peaks should be made very carefully, because the intensity is always related to the measurement conditions.

10.2. TEM

Electron microscopy is one of the most important modern techniques for characterizing the geometry, periodicity, pore morphology, etc. of mesoporous materials [145, 146]. In addition, HRTEM can also provide information about superlattices, defects, twin growth, and breaking edge analysis for mesoporous materials.

Analysis of TEM images is always complicated. Artificial images and incorrect indexation often take place. To obtain more information on mesoporous materials, it is necessary to combine TEM results with electron diffraction analysis and XRD patterns. It should be noted that TEM data provide only local periodicity and structure information in small domains and cannot give us information for whole samples, whereas XRD can give information for whole bulk samples. TEM can be used to analyze the structure of mesoporous materials. Terasaki has made an excellent contribution in this field [83, 145, 146]. He and co-workers obtained high-quality images of mesoporous materials by defocus TEM, combined with XRD and N_2 sorption isotherm results; in this way they creatively established the energy distribution model. With this model they effectively resolved several mesoporous material structures, such as SBA-1, SBA-6, SBA-16, etc. (Fig. 10). Recently they found a novel method for resolving the structure of molecular sieves by TEM, which was expected to provide more useful tools for structural characterization in the field of molecular sieves [146].

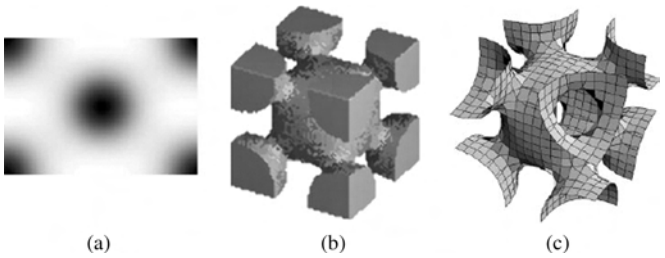


Figure 10. Direct image of 3D pore structure of SBA-16. (a) Electrostatic potential map of SBA-16 parallel to (110) through the centre of the cell. (b) 3D arrangement of a cavity and its interconnection. Black corresponds to the cavity. (c) Mean cavity surface for SBA-16. Reprinted with permission from [145], Y. H. Sakamoto et al., *Nature* 408, 449 (2000). © 2000, Macmillan Magazines Ltd.

10.3. Physical Adsorption

Physical adsorption is also one of the most effective modern tools for the characterization of the pore structure of mesoporous materials. Based on the shape of the isotherms and hysteresis loop, pore structures, including micro-, meso-, macropore, cylindrical pore, bottle-like pore, and so on, can be obtained. Pore size distribution, pore volume, and surface area can also be calculated. The probes used are usually Ar and N_2 molecules. Because the atom diameter of Ar is smaller than that of N_2, the measurement error for the pore size from Ar is smaller [12, 13]. Unfortunately, liquid Ar is not conveniently available. N_2 sorption measurements are more convenient to carry out. A typical curve for mesoporous materials is the IV type sorption isotherm. The steeper the adsorption volume change is, the more uniform the pore size distribution is. Mesopore size is generally calculated from the adsorption branches of the isotherms, and the window size of the pore is calculated from the desorption branches. Typical N_2 sorption isotherms of mesoporous silica materials are shown in Figure 11. Pore structure can be roughly judged from the shape of the hysteresis loop. MCM-41 or SBA-15 has cylindrical pores, and therefore their window sizes are the same as their mesopore diameters. The hysteresis loop is parallel and belongs to the H_1 type; therefore, the pore diameters calculated from desorption branches are the same as those from adsorption branches (Fig. 11). Cubic mesoporous materials with caged structures have small windows (smaller than 4 nm) [145]; therefore an H_2 type hysteresis loop is observed. Because of the physicochemical properties of N_2, the hysteresis loop always ends near the point of $P/P_0 \sim 0.4$. Several models, such as BJH (Barrett-Joyner-Halenda), HK (Horvath-Kawazoe), and BdB (Broekhoff-de-Boer), can be used to calculate the pore sizes and their distribution. The BJH model is mostly utilized by mesoporous material researchers. For a simple comparison, the BJH model is advisable, but some problems have been observed when this model has been used to calculate the pore sizes. The pore diameters calculated from the adsorption branches for cylinder pore MCM-41 or SBA-15 are not the same as those from the desorption branches, and the 4V/A value is not equal to the radius of the cylinder. It was found recently that the pore size calculated from the BdB model is more exact and is in agreement with that from other characterization

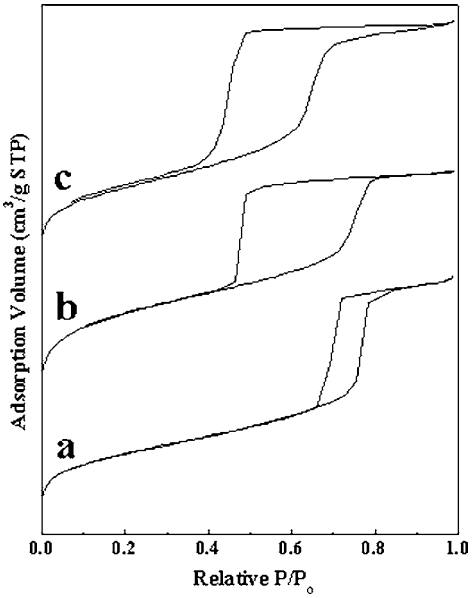


Figure 11. Typical nitrogen sorption isotherms of mesoporous materials with different pore structures. (a) SBA-15. (b) SBA-16. (c) SBA-1.

results [149]. In the past, because there have been no reference samples for the measurements of mesopore sizes, adsorption models and theories have been difficult to prove. The appearance of periodic ordered mesoporous materials such as SBA-15 and MCM-41 has made it possible to further develop the adsorption theory. Theoretically considered, pore diameters smaller than 300 nm can be measured by N_2 isotherms, but the practical error is large when the pore diameter is larger than 100 nm or even 50 nm. Pore diameters larger than 100 nm are normally measured by mercury extrusion. In fact, other adsorption probe molecules with different dynamic diameters, such as water, alcohol, *n*-hexane, cyclohexane, and aromatic hydrocarbons, can also be used in the measurement of pore diameters, which can provide an applicable pore diameter.

10.4. Solid-State NMR

Solid-state NMR is an important technique as well for learning about the framework structures of mesoporous materials [99]. Because the walls of mesoporous silica materials are amorphous, typical ^{29}Si-NMR spectra show fairly broad peaks, which include Q^2, [$=$Si(OH)$_2$], Q^3, [$\equiv$Si—OH], Q^4, [Si(—O)$_4$], at the chemical shift between -90 and -110 ppm. The quantity of hydroxy groups on the frameworks of mesoporous materials can be evaluated from the ratio of $Q^2/Q^3/Q^4$. Generally, the smaller the quantity is, the higher the thermal stability and hydrothermal stability are. Sometimes the evaluation is not correct. For example, compared with calcined MCM-41, calcined SBA-15 has a larger Q^3/Q^4 ratio because abundant micropores are located on the silica walls and result in a larger number of hydroxy groups, but the hydrothermal stability (heating in boiling water) of calcined SBA-15 is higher than that of MCM-41.

^{27}Al-NMR, ^{1}H-NMR, and ^{13}C-NMR [99] can also be applied to the characterization of mesoporous materials.

It has been known that the resonance peak for the tetrahedral Al species appears at 50 ppm and that for octahedral Al near 0 ppm. Though it cannot be concluded that octahedral Al species are not located on the inorganic framework (because the inorganic walls of mesoporous materials are amorphous), it is reasonable to believe that the octahedral Al species are located on the surface of the walls of the mesoporous materials. The formation mechanism of mesostructures and the interaction of organic species (surfactants) with inorganic species can also be studied by 2D-NMR [99].

11. MORPHOLOGY

Mesoporous materials can be synthesized in either basic or acidic conditions. In basic conditions, the Coulombic interaction between organic templates and inorganic precursors is relatively strong; furthermore, the silica walls are robust, therefore the morphology of the resulted mesoporous materials is quite rigid and in some cases sharp polyhedrons are obtained [82]. Under acidic conditions, hydrogen bonding is relatively weak, and the hydrolysis and condensation of silica species can be controlled efficiently by adjusting the temperature, pH, ionic strength, etc. In this way, combined with emulsion chemistry, mesoporous materials with fiber, film, monolith, and sphere morphologies, even mesoporous silica single crystals, have been obtained.

11.1. Fibers

Mesoporous fibers can be prepared by two methods. One synthesis approach is growth in a solution [150–152]. The synthesis of mesoporous silica fibers has been carried out at the interface of oil (long alkyl chain alcohols or hexane) and aqueous solutions (i.e., CTAB acidic solutions) when TEOS was used as a silica source [151]. The oil is not necessary when tetrabutylorthosilicate (TBOS) is used as a silica source. Under static conditions, mesoporous silica fibers can be obtained from the aqueous phase that are 1–10 μm in diameter and 100 μm to 5 cm in length (Fig. 12) [150, 151].

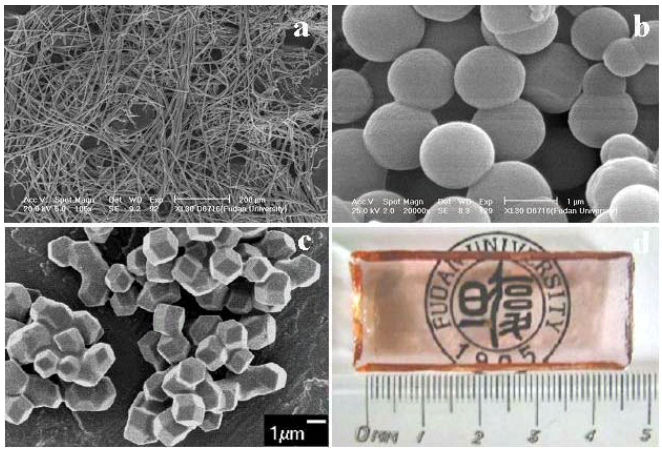


Figure 12. Selective SEM images of mesoporous materials synthesized with different morphologies. (a) Fibers. (b) Spheres. (c) Single crystals. (d) Monoliths.

For the pore structure of hexagonal mesoporous silica fibers synthesized by this method, it was once suspected that the hexagonally arrayed channels are parallel to the long axis of the fibers. Recent results suggest that the mesopores in such fibers run in a circular direction around the axis. The solution growth method has been successfully applied to the synthesis of SBA-15 fibers, with triblock copolymer P123 as a template and TMOS (tetramethoxysilane) as a silica source [93, 152]. The fast precipitation of products together with stirring is essential to the formation of mesoporous fibers [126]. Moreover, the mesoporous channels of the SBA-15 fibers obtained proved to be parallel to the long axis. Another approach to synthesizing mesoporous fibers was simply performed by drawing fibers from concentrated gel solutions [153]. These fibers do not have a well-defined mesopore orientation because the formation is too fast.

11.2. Films

Continuous and crack-free films with accessible mesopores may offer important potential for use in separation, adsorption, and electronic/optical applications. Generally dip-coating or spin-coating methods were utilized to synthesize such mesoporous films by an evaporation-induced self-assembly process [91–93]. The substrate at which the film is grown is an important factor. Yang et al. first reported the synthesis of a hexagonally ordered mesoporous thin film at a cleaved mica substrate with the use of an acidic aqueous mixture of CTAC (cetyltrimethyl ammonium chloride) surfactant and TEOS; TEM revealed that the pore channels were parallel to the mica surface [94]. A later report by Yang et al. demonstrated that ordered mesoporous silica films could be formed at the interface of air and water without a solid substrate [95]. Other research revealed that mesoporous thin films could be grown at a hydrophobic interface such as carbon, instead of commonly used hydrophilic substrates such as mica and glass slides [154].

Similar to what has been observed in the synthesis of powder samples, the choice of the structure-directing agents may result in the variation of the structure and pore size of mesoporous films. Stucky and co-workers reported the synthesis of 3D hexagonal mesoporous thin films with the use of Gemini surfactants as the templates, in which the *c* axis of the oriented film is perpendicular to the substrates [93, 96]. Using the block copolymers as the templates, the same group reported that highly ordered, large-pore, 2D and 3D mesoporous silica and transitional metal oxide films can be synthesized [90, 92]. By changing the silica source, Brinker and co-workers reported the synthesis of highly ordered hybrid mesoporous films [97].

11.3. Monoliths

Mesoporous silica monoliths with enhanced elasticity and ductility could be synthesized with block copolymers as the templates in a nonaqueous solution based on the sol–gel process [87]. A representative SEM image for mesoporous silica monolith templated by triblock copolymer is shown in Figure 12. Such rigid and transparent monoliths have potential uses in optics. Organic molecules such as butanol, pentanol, hexanol, or octane can be used as the cosurfactants

to control the mesostructure and the pore size of the final monolith products [98]. Zhao et al. synthesized 3D sponge-like macro- and mesoporous silica monoliths by employing the phase separation of concentrated salt solutions in the presence of block copolymer surfactants [93]. The hierarchical pore structures may be important for mass transportations in catalysis and separation. It is noted that cationic surfactants such as CTAB can also be used in the synthesis of silica monoliths.

Normally, the formation of crack-free mesoporous monoliths is too slow; it takes several days to a month [99, 100]. To solve this problem, recently we demonstrated a method in which liquid paraffin medium protected against solvent evaporation to quickly yield transparent, crack-free, large, highly ordered mesoporous monoliths. This method is quite simple; it just includes covering an inert layer liquid of paraffin as the morphology "protector," to reduce the damage caused by the internal pressure gradient. By accelerating the evaporation process of a silica gel by inert liquid paraffin with heating, the processing period is shortened to within 8 h [101].

11.4. Spheres

With a similar method for synthesizing mesoporous powders, mesoporous silica spheres with different diameters (Fig. 12) can be formed under either basic or acidic media [97, 102, 116–120, 155–157]. The key factors are stirring rate, silica source, and basic source. Marble-like mesoporous silica hard spheres ~1 mm in diameter can be prepared by using CTAB as a template [116]. It is essential for successful synthesis to use TBOS as a silica source. The stirring rate can be used to control the sizes of such spheres. These mesostructured silica spheres may be readily applied as catalysts when metal ions such as Cu^{2+} are doped into the frameworks. Under acidic conditions, Schacht et al. synthesized hollow spheres with 2D hexagonal and 3D hexagonal mesostructured shells, with the use of emulsion templating [117]. With ammonia as the base, uniform mesoporous silica spheres with a diameter size of 1–10 μm can be prepared with CTAB as a template [118, 155–157]. In the presence of dimethyl formamide (DMF) as a cosolvent and CTAB as the structure-directing agent, Luo et al. synthesized mesoporous spheres with a controllable nanometer scale (80–800 nm) [118]. A fast aerosol method has been developed by Brinker et al. to synthesize mesostructured spheres with hexagonal, cubic, and vesicle structures; the whole process can be completed within 6 s [102, 119]. Zhao et al. have synthesized hollow spheres with ultralarge mesopore (50 nm) wall structures by reverse emulsion templating [120]. Millimeter-size SBA-16 spheres have been synthesized in the same group with the use of block copolymers as the templates in the presence of inorganic salts such as NaCl. The salts can improve the structural regularity of mesoporous spheres [158].

11.5. Mesoporous Crystals and Polyhedrons

Mesoporous materials possess no short-range order on the atomic scale but obvious order on the mesoscopic scale. Unlike microporous materials that usually show well-defined crystal morphology in accord with their atomic order, it is unusual to synthesize mesoporous materials with crystal shapes [159]. A general method for growing crystal-like mesoporous materials has not been reported yet. The static condition is one of the keys to yields of crystal-like materials [151]. Kim and Ryoo first reported the synthesis of MCM-48 crystals (*Ia3d*) with a truncated rhombic dodecahedral shape [160]. Guan and Sayari et al. have reported the synthesis of hybrid cubic mesoporous crystals (*Pm3n*) with a well-defined decaoctahedron shape [161]. SBA-1 (*Pm3n*) has also been synthesized with crystal morphologies with a large number of facets [114]. In the above cases, ionic surfactants have been utilized to synthesize mesoporous crystals. An exception has been reported by Zhao and co-workers; cubic mesoporous silica single crystals (Fig. 12) have been synthesized with the use of a commercial non-ionic block copolymer as a template under acidic conditions [68]. This is achieved by using inorganic salts to increase the interaction between silicate species and non-ionic block copolymers. These single crystals possess exclusively uniform rhombdodecahedron shapes (~1 μm) with ~100% crystal yield and uniform large pores (up to 7.4 nm). The mesopore array in each crystal face is resolved by TEM, further confirming that these crystals are perfect single crystals with a body-centered cubic space group (*Im3m*) [68].

11.6. Other Morphologies

By careful choice of the pH of the reacting systems, MCM-41 materials have been synthesized with a tubules within tubules (TWT) morphology [162, 163]. With neutral double-head amines, $C_nH_{2n+1}NH(CH_2)_2NH_2$, as the templates, mesostructured silica vesicles can be obtained that are hydrothermally ultrastable (the structure is retained in 100 °C water for 150 h) [164].

12. MECHANICAL PROPERTIES AND APPLICATIONS

12.1. Mechanical Properties

Mechanical stability is an important property of nanoporous materials, which normally act as adsorbents or catalysts for most practical applications. Before the processes of adsorption and catalysis reaction, the fine powders obtained are routinely compacted at high pressure into pellets. High mechanical stability is a critical requirement for conserving the pore volume, surface area, and pore diameter of the original material. Studies of the mechanical properties of mesoporous materials have been carried out by several groups, especially on some important materials such as MCM-41, MCM-48, and SBA-15 [165–171]. The experiments were performed by treating the mesoporous samples with different external pressures and then characterizing their structure by XRD or NMR and adsorption properties by N_2, subcritical organic agents (benzene, *n*-heptane, or cyclohexane), or mercury porosity [165–171]. Broad X-ray diffraction peaks and weak adsorption capacities are always observed for the mesoporous materials after they have been compressed [165–171]. For MCM-41, the

BET surface area and pore volume decreased considerably with increasing pressure to 86 MPa, and the material was essentially destroyed at 224 MPa [165]. MCM-48 possessed better mechanical stability than MCM-41, possibly because of the 3D structure. Its adsorption capacity was affected by mechanical compression between 20 and 400 MPa before the material was finally destroyed at a high pressure (480–600 MPa) [167]. However, other research results have suggested that there are no remarkable changes in the mechanical stability of mesoporous materials such as MCM-41, MCM-48, HMS, FSM-16, KIT-1, and SBA-15 at relatively low pressures; all materials collapsed at a maximum pelletizing pressure of 450 MPa [169].

12.2. Catalysis

Zeolites have led to a brand new concept of "size- and shape-selective catalysis" in the petroleum industry; however, the pore size of zeolites is too small to process large molecules such as heavy oil. Mesoporous materials with a large pore size may favor the diffusion rate for large molecules and decrease the second cracking; therefore, the catalytic properties of mesoporous aluminosilcates in cracking and the hydrogenation process have been studied extensively.

In addition to the application of petroleum cracking, mesoporous materials showed excellent catalytic activity in many typical reactions, such as oxidation, acid catalysis, and hydroxylation. Bekkum et al. used nanoporous materials (HMCM-41) as catalysts in acid catalytic reactions such as acylation of 1-naphthol [172, 173]. It was found that HMCM-41 shows a much higher catalytic activity and selectivity (up to 99%) compared with HY zeolites. Sodium ion- or cesium ion-exchanged MCM-41 showed a high selectivity and catalytic activity for Knoevenagel and aldol condensation reactions, and cesium-doped MCM-41 was a good basic catalyst and displayed high activity in the Michael addition [174]. Since mesoporous aluminosilicates show mild surface acidity, improvement of the surface acidity of mesoporous materials by loading is necessary for the application of acidic catalytic reactions. Kozhevnikov et al. loaded heteropolyoxometallate acid, $H_3PW_{12}O_{40}$ (PW), into MCM-41 materials to obtain a PW/MCM-41 catalyst that exhibited shape selectivity and higher reaction activity than H_2SO_4 or bulk PW [175]. Corma and co-workers prepared NiMo/MCM-41 catalysts [12 wt% MoO_3, 3 wt% NiO] and found that such materials showed much higher activities for the hydrocracking of vacuum gas oil compared with amorphous silica-alumina and USY zeolites [176]. The large pore size and high surface area of mesoporous MCM-41 materials are important in the above-mentioned catalytic reactions.

Nanoporous materials can also be used as supports in the famous Ziegler-Natta polymerization. Woo and O'Hare showed that the polymers obtained have a high density and a large particle size [177, 178]. In 1999, Kageyama et al. synthesized 2D hexagonal mesoporous fibers and used them as a catalyst for the polymerization of ethylene [179]. The products are polyethylene nanofibers with diameters between 30 and 50 nm and an average molecular weight over 6,000,000, which is ten times larger than that of conventional polyethylene. It is believed that this polymerization was carried out in 1D mesoporous channels, and the morphology of the polyethylene is very dependent on the shape of the channels of supporting mesoporous materials. This achievement has led to the new idea that researchers might develop a new catalytic process for nanoporous materials.

12.3. Optics

Nanoporous materials have long, straight nanochannels and may be used as waveguides in optics. Marlow et al. have coated mesoporous fibers with laser dye molecules and found that the fibers obtained can emit light parallel to their longitudinal axis with amplified intensity [180]. The results imply that the fibers have resonant structure and can be used as laser materials such as optical waveguides. Yang et al. have prepared patterned mesoporous oxide materials with a low refractive index by combining micromolding, polystyrene sphere templating, and cooperative assembly of inorganic sol–gel species with amphiphilic triblock copolymers [181]. The resulting materials show hierarchical ordering over several discrete and tunable length scales ranging from 10 nm to several micrometers, which may be useful for constructing an optic loop. Wirnsberger et al. also coated optical fibers with mesostructured silicate film containing laser dye molecules. Emission with a low valve value and signal amplification has been observed when the functionalized fibers were activated by laser [182]. It is expected that the new composite can be used as a sensor to detect the gas molecules that are absorbed in the mesopore channels. Dag et al. have prepared silicon nanoclusters with a size of about 1 nm on mesoporous films by low-temperature CVD [183], which possessed luminescent life at about the microsecond level, much less than that of general porous silicon. Researchers hoped that this material could be used in silicon-based diodes. More recently, Karen et al. synthesized Eu^{3+}-doped mesoporous TiO_2 films with cubic structures and found that the films were intensively luminescent [184].

12.4. Nanoreactors

A challenge in material science is to control the size and shape of compounds at the nanometer scale. It is certain that mesoporous materials with ordered pore structures and defined pore connectivity can be utilized as ideal nanoreactors to confine the growth of nanomaterials. Yang and co-workers prepared uniform Ag nanorods with 5–6-nm diameters by the pyrolysis of $AgNO_3$ in the 1D ordered mesopore channels of SBA-15 [185]. In work similar to this, Stucky and Ryoo have used mesoporous silica SBA-15 and MCM-41 as hard templates to prepare metal Pt, Ag, and Au nanorods [186–188]. Mesoporous silica FSM-16 [189] and HMS [190] can also be used as hard templates to prepare metal Ag, Rh, and alloy Pt-Rh nanorods. Moreover, with this hard template approach, binary sulfides such as CdS and PbS have been used to fill in the nanoreactors of mesoporous materials, which resulted in the formation of uniform CdS and PbS nanoparticles and nanorods [191, 192].

12.5. Other Applications

Researchers found that mesoporous silica materials functionalized with MPTS can selectively absorb heavy metal ions such as Hg^{2+}. Instead the absorption amount of Cd^{2+}, Pb^{2+}, and Zn^{2+} ions etc. is quite low [193, 194]. The results suggest that these functionalized mesoporous materials are expected to be used in the treatment of wastewater in the nuclear industry. Dai et al. have grafted multiorganic functional groups to mesoporous SiO_2 materials and found that the resulting materials can be used as an adsorbent to selectively separate some heavy metal ions [195]. Such mesoporous materials are also attractive candidates for hosting large molecules, including proteins. When the pore size was varied and the surface silanol groups were derivatized, proteins of different sizes with mesoporous silicate materials were specifically selected [196]. The sequestration of anionic proteins (conalbumin, M_W 77,000; chicken egg ovalbumin, M_W 44,000; and soybean trypsin inhibitor protein, M_W 14,000) from aqueous solutions by 3-aminopropanetri(ethoxy)silane (APTS) functionalized SBA-15 and MCF, as well as the release of the sequestered proteins with an increase in the ionic strength, has been reported [196]. In such a procedure, both size exclusion and ion-exchange chromatography techniques are employed, suggesting the potential application of the nanoporous materials as chromatography in a protein separation. With the modification of SBA-15 with dimethyloctadecylchlorosilane, the nanoporous material has been successfully used as a high-performance liquid chromatography substrate to separate biomolecules, including peptides and proteins (lysozyme, bovine serumin, myoglobin, and ovalbumin).

13. MACROPOROUS MATERIALS

Ordered macroporous materials with a pore size in the range of the wavelength of light have unique optical properties and are the best photonic crystals [197, 198]. Unfortunately, a lot of problems are crearted by using colloidal microspheres as the templates to prepare this kind of photonic crystal. The sizes of the macroporous materials are still not large enough to serve as a photonic crystal. Normally, three methods can be used to synthesize ordered macroporous materials:

(1) The sol–gel templating method. With the use of uniform-sized polystyrene spheres (PS spheres) as the templates, and the impregnation of inorganic sol into the voids between the close-packed PS spheres, after a sol–gel process and calcinations for removing templates, ordered macroporous materials with a 3D spherical pore (large cage) structure can be obtained (Fig. 13). The pore size can be tuned at the range of 50–500 nm by varying the diameter of the PS spheres. With this method, macroporous SiO_2 [199], TiO_2 [200], ZrO_2, and Al_2O_3 [201] materials have been synthesized. These macroporous materials have uniform pore diameters and normally a face-centered cubic (fcc) or hexagonal close packing (hcp) pore structure. The surface of emulsified PS microspheres (diameter in the range of 50–500 nm) can be modified to negative charges (sulfates) or positive charges (amines) [202]. In this way, the inorganic sols will easily to fill up the voids of charged and closely packed microspheres. The key point of the technology is to pack the ordered arrays of the colloidal microspheres. Closely packed PS spheres can be achieved by evacuation, sedimentation, capillary extrusion, and flat pattern. As well, SiO_2 colloidal spheres can also be used as the templates for synthesizing ordered macroporous materials. Compared with PS spheres, SiO_2 colloidal spheres have better thermal stability, and, combined with the chemical vapor deposition (CVD) technique, they can be used to prepare macroporous materials with a high melting point, such as C [203], Si [204], SiC [205], and so on. In addition, silica spheres can also be used in the synthesis of macroporous materials with organic macromolecular networks. With the combination of electrodeposition, metal, metal alloy, semiconductor sulfides, and selenite macroporous materials have been successfully synthesized (Fig. 13) [206]. In addition, the zeolite nanoparticles can also be used as the precursors for synthesizing macroporous materials with zeolite crystal frameworks. Meso-macro hierarchical ordered porous materials can be prepared as well.

(2) Solid template method. Macroporous silica materials can be orientationally mineralized on bacteria thread [207]. With this method, porous meso-macro materials can also be obtained. The macropores are paralleled long channels, pore diameters are in the range of micrometers, and the wall thickness is 50–200 nm.

(3) Using surfactant-oil emulsion as the template. In the sol–gel process, with the emulsion formed by a surfactant with oil or water as a template, an ordered macroporous material can be synthesized by deposition of inorganic oxides on the external surface of the emulsion droplets. The pore diameter is in the range of 50 nm to several micrometers. The emulsion drops produced by oil and formamine have uniform sizes, which are suitable for use as the templates and can be stabilized by PEO-PPO-PEO block copolymer surfactants. Ordered macroporous SiO_2, TiO_2,

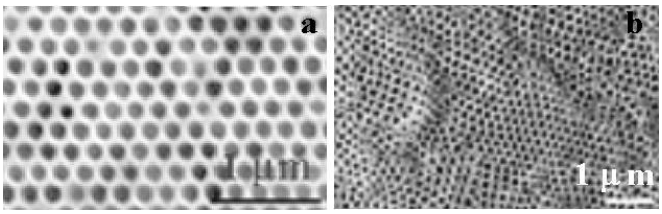


Figure 13. SEM image for (a) ordered macroporous alloy CoSn materials prepared by electrodeposition, with PS spheres as templates; (b) 3D ordered macroporous silica prepared with latex particles 300 nm in diameter as templates. Reprinted with permission from [198], O. D. Velev et al., *Nature* 389, 447 (1997). © 1997, Macmillan Magazines Ltd.

Table 1. List of mesoporous materials and their properties.

Mesoporous materials	Structure director or template	Space group	Pore size (nm)	Ref.
MCM-41	C_nTMA^+ ($n = 12$–18)	*p6mm*	2–5	[12, 13]
MCM-48	C_nTMA^+ (C_nTAB, C_nTACl, C_nTAOH, $n = 14$–20)	*Ia3d*	2–5	[12, 13]
MCM-50	C_nTMA^+ ($n = 12$–18)	Lamellar	2–5	[12, 13]
SBA-1	$C_{16}H_{33}N(CH_3)(C_2H_5)^{2+}$-$C_{16}H_{33}N(C_2H_5)^{3+}$	*Pm3n*	2–4	[24, 27, 52]
SBA-2	$C_{12\text{-}3\text{-}1}$, $C_{14\text{-}3\text{-}1}$, $C_{16\text{-}3\text{-}1}$, $C_{18\text{-}3\text{-}1}$, $C_{18\text{-}6\text{-}1}$, $C_{20\text{-}3\text{-}1}$	$P6_3/mmc$	2–4	[24, 27, 52]
SBA-3	CTAB	*p6mm*	2–4	[52]
SBA-6	$C_{18\text{-}3\text{-}1}$	*Pm3n*	2–3	[24, 27]
SBA-11	Brij 56 ($C_{16}EO_{10}$)	*Pm3m*	2.5	[63]
SBA-12	Brij 76 ($C_{18}EO_{10}$)	$P6_3/mmc$	3.1	[63]
SBA-14	$C_{12}EO_2$	*cubic*	2.2	[63]
SBA-15	P123	*p6mm*	5–15	[62, 63]
SBA-16	F127	*Im3m*	5.4	[63]
FDU-2	$[C_mH_{2m+1}N^+(CH_3)_2CH_2CH_2N^+(CH_3)$-$2CH_2CH_2CH_2N^+(CH_3)_3 \cdot 3Br^-]$ ($C_{m\text{-}2\text{-}3\text{-}1}$, $m = 14, 16, 18$)	*Fm3m*	2–3	[209]
FDU-5	P123	*Ia3d*	5–10	[89]
MSU-V	$H_2N(CH_2)_nNH_2$ (bolaamphiphiles $n = 12$–22)	Lamellar	2.0–2.7	[210, 211]
MSU-G	$C_nH_{2n+1}NH(CH_2)_2NH_2$	Vesicle-like	2–4	[164]
MSU-H	P123	Hexagonal	7.6–11.9	[212]
MSU-X	Polyethylene oxide (PEG)	Wormhole	2.0–5.8	[40]
HMS	Dodecylamine	Wormhole	1.9–4.3	[39, 41]
CMK-1	MCM-48	*I432*	3	[75]
CMK-3	SBA-15	*p6mm*	4.5	[76]
CMK-4	MCM-48	*Ia3d*	3	[82]
CMK-5	SBA-15	Hexagonal	2 and 4[a]	[77]

[a]The material has two sets of pores.

and ZrO_2 materials have been successfully synthesized by this method [208]. With the use of the liquid drops of some inorganic salts like NaCl, KCl, and NH_4Cl, meso-macroporous silica membranes can be prepared [67]. Their macro-pores are not uniformly distributed and can be tunable in the range of 500 nm to 5 μm, but these meso-macroporous silica membranes show great potential applications in catalysis and separation [67].

14. CONCLUSIONS

The surfactant templating approach can be used to prepare highly ordered nanoporous materials with variable compositions. Nanoporous materials have long-range periodic pore 2D or/and 3D structures at the nanoscale, narrow pore size distribution, and extra large surface areas. Despite the great advances that have been achieved in nanoporous materials, some challenges remain to be further explored. From the point view of synthesis and future applications, it is very important to synthesize ordered nanoporous materials with crystalline frameworks, especially if the pore system is more than 1D and the structure exhibits thermal and hydrothermal stability. For energetic considerations [213], the thermodynamic barriers in the assembly process contribute little to the difference in frameworks between traditional crystalline zeolites and amorphous mesoporous materials, hence the dynamic behavior of formation of mesoporous materials and the limitation of the surfactant assembly pathway. For example, the hydrothermal temperature in the synthesis of mesoporous materials is normally lower than that in the synthesis of microporous materials; the surface charge density of structure-directing agents in the former case is usually much lower than in the latter case.

Another challenge facing the nanoporous materials community is to step from laboratory studies to industrial applications. This is even more important for the continuing study of mesoporous materials. It has been ten years since the report of mesoporous materials by Mobil scientists, and the application potential of such materials has been widely studied in catalysis, separation, and optics, and for nanoreactors, low dielectric constant materials, and sensors. Forthcoming commercialized mesoporous materials and related applications (1, see http://www.schumacher.com/products/DESCRIP/CHEM/Meso.htm for details; 2, see [214]) will stimulate more research interest in this area, and much more exciting developments in nanoporous materials will be delivered.

GLOSSARY

Hydrothermal synthesis Synthesis processes under a certain pressure (1–100 MPa) and a certain temperature (100–1000 °C), utilizing the chemical reaction in the liquid.

M41S A family of ordered mesoporous aluminosilicates discovered by Mobil researchers in 1992.

Macroporous material Porous materials with a pore diameter larger than 50 nm.

Mesoporous material Porous materials with a pore diameter in the range of 2–50 nm.

Microporous material Porous materials with a pore diameter smaller than 2 nm.

Nanoporous material Porous materials with a pore diameter at the nanometric scale.

Self-assembly A coordinated action of independent entities under distributed (i.e., noncentral) control to produce a larger structure or to achieve a desired group effect. It is a method of constructing devices and machines whose dimensions are measured in nanometers (units of 10^{-9} m or millionths of a millimeter). The technique works by simulating the way in which biological systems build molecules, viruses, cells, plants, and animals.

Sol–gel process A versatile solution process for making ceramic and glass materials involving the transition of a system from a liquid "sol" (mostly colloidal) into a solid "gel" phase by hydrolysis and polymerization of molecules or ions.

Structure-directing agent Structure-directing additives, often organics or polymers, in addition to the constituent-building units, which may be inorganic compounds. This template can serve to fill space, balance charge, and direct the formation of a specific framework. In this early context, "templating" and "structure-directing" were synonymous.

Surfactant Also known as a surface-active agent, a soluble compound that reduces the surface tension of liquids or reduces interfacial tension between two liquids or a liquid and a solid.

ACKNOWLEDGMENTS

We acknowledge the financial supports by the National Science Foundations of China, State key research program, Shanghai Science and Technology Committee, Ministry of Education of Chinese. We are very grateful for useful discussions with Dr. Q. Huo, B. Tian, X. Liu, J. Fan, F. Zhao, etc.

REFERENCES

1. H. V. Bekkum, E. M. Flanigen, and J. C. Jansen, "Introduction to Zeolite Science and Practice." Elsevier, Amsterdam, 1991.
2. R. Szostak, "Molecular Sieves: Principles of Synthesis, and Identification." Blackie Academic and Professional, London, 1998.
3. IUPAC Manual of Symbols and Terminology, *Pure Appl. Chem.* 31, 578 (1972).
4. J. Y. Ying, C. P. Mehnert, and M. S. Wong, *Angew. Chem. Int. Ed. Engl.* 38, 56.
5. A. Corma, *Chem. Rev.* 97, 2373 (1997).
6. G. W. Brindley and R. E. Semples, *Clay Miner.* 12, 229 (1997).
7. N. Lahav et al., *Clays Clay Miner.* 26, 107 (1978).
8. S. Yamanaka and G. W. Brindley, *Clays Clay Miner.* 26, 21 (1978).
9. S. Yamanaka and G. W. Brindley, *Clays Clay Miner.* 27, 119 (1979).
10. T. J. Pinnavaia, *Science* 220, 365 (1983).
11. M. L. Occelli and R. M. Tindwa, *Clays Clay Miner.* 31, 22 (1983).
12. C. T. Kresge, M. E. Leonowicz, W. J. Roth, J. C. Vartulli, and J. S. Beck, *Nature* 359, 710 (1992).
13. J. S. Beck, J. C. Vartuli, W. J. Roth, M. E. Leonowicz, C. T. Kresge, K. D. Schmitt, C. T. W. Chu, D. H. Olson, E. W. Sheppard, S. B. McCullen, J. B. Higgins, J. L. Schlen, and J. E. ker, *J. Am. Chem. Soc.* 114, 10834 (1992).
14. F. DiRenzo, H. Cambon, and R. Dutartre, *Microporous Mater.* 10, 283 (1997).
15. 3V. Chiola Ritsko and C. D. Vanderpool, U.S. Patent 3,556,725, 1971.
16. T. Yanagisawa, T. Shimizu, K. Kuroda, and D. Kato, *Bull. Chem. Soc. Jpn.* 63, 1988 (1990).
17. S. Inagaki, Y. Fukushima, and K. Kuroda, *Chem. Commun.* 8, 680 (1993).
18. M. E. Leonowicz, J. A. Lawton, S. L. Lawton, and M. K. Rubin, *Science* 264, 1910 (1994).
19. G. J. Kennedy, S. L. Lawton, and M. K. Rubin, *J. Am. Chem. Soc.* 116, 11000 (1994).
20. M. K. Rubin and P. Chu, U.S. Patent 4,954,325, 1990.
21. J. C. Vartuli, K. D. Schimitt, and C. T. Kresge, *Stud. Surf. Sci. Catal.* 84, 53 (1994).
22. J. C. Vartuli, C. T. Kresge, and M. E. Leonowicz, *Chem. Mater.* 6, 2070 (1994).
23. R. J. Agrauer, U.S. Patent 3,702,886, 1972.
24. Q. Huo, D. I. Margolese, U. Ciesla, D. G. Demuth, P. Feng Gier, T. E. Sieger, P. A. Firouzi, B. F. Chmelka, F. Schüth, and G. D. Stucky, *Chem. Mater.* 6, 1176 (1994).
25. C. Y. Chen, S. L. Burkett, H. X. Li, and M. E. Davis, *Microporous Mater.* 2, 27 (1993).
26. A. Monnier, F. Schuth, Q. Huo, D. Margolese, R. S. Maxwell, G. D. Stucky, M. Krishamurty, P. Petroff, A. Firouzi, M. Janicke, and B. F. Chmelka, *Science* 261, 1299 (1993).
27. Q. Huo, D. I. Margolese, U. Ciesla, P. Y. Feng, T. E. Gier, P. Sieger, R. Leon, P. M. Petroff, F. Schuth, and G. D. Stucky, *Nature* 368, 317 (1994).
28. Q. Huo, R. Leon, P. M. Petroff, and G. D. Stucky, *Science* 268, 1324 (1995).
29. G. D. Stucky, Q. Huo, A. Firouzi, B. F. Chemalka, S. Schacht, I. G. Voigt-Martin, and F. Schuth, *Stud. Surf. Sci. Catal.* 105, 3 (1997).
30. A. Firouzi, D. Kumar, L. M. Bull, T. Besier, P. Sieger, Q. Huo, S. A. Walker, J. A. Zasadzinski, C. Glinka, J. Nicol, D. Margolese, G. D. Stucky, and B. F. Chmelka, *Science* 267, 1138 (1995).
31. Q. Huo, D. Z. Hao, J. Feng, K. Weston, S. K. Buratto, G. D. Stucky, S. Schacht, and F. Schüth, *Adv. Mater.* 9, 974 (1997).
32. F. Marlow, B. Spliethoff, B. Tesche, and D. Zhao, *Adv. Mater.* 12, 275 (2000).
33. Z. Y. Yuan and W. Z. Zhou, *Chem. Phys. Lett.* 333, 427 (2001).
34. Q. Lu, F. Gao, and D. Zhao, *Microporous Mesoporous Mater.* 56, 219 (2002).
35. S. Inagaki, Y. Fukushima, and K. Kuroda, *Stud. Surf. Sci. Catal.* 84, 125 (1994).
36. C. Y. Chen, S. Q. Xiao, and M. E. Davis, *Microporous Mater.* 4, 1 (1995).
37. B. T. Holland, P. K. Isbester, C. F. Blanford, E. J. Munson, and A. Stein, *J. Am. Chem. Soc.* 119, 6796 (1997).
38. M. Yada, H. Takenaka, and T. Kijima, *J. Chem. Soc., Dalton Trans.* 1547 (1998).
39. P. T. Tanev and T. J. Pinnavaia, *Chem. Mater.* 8, 2068 (1996).
40. S. A. Bagshaw, E. Prouzet, and T. J. Pinnavaia, *Science* 269, 1242 (1995).
41. P. T. Tanev and T. J. Pinnavaia, *Science* 267, 865 (1995).
42. S. A. Bagshaw and T. J. Pinnavaia, *Angew. Chem. Int. Ed. Engl.* 35, 1102 (1996).
43. E. Prouzet, F. Cot, G. Nabias, A. Larbot, P. Kooyman, and T. J. Pannavaia, *Chem. Mater.* 11, 1498 (1999).
44. D. M. Antonelli and J. Y. Ying, *Angew. Chem. Int. Ed. Engl.* 35, 426 (1996).
45. C. J. Brinker and G. W. Scherer, *Sol–Gel Science.* Academic, New York, 1990.

46. C. F. Cheng, H. He, W. Zhou, and J. Klinowski, *Chem. Phys. Lett.* 244, 117 (1995).
47. Z. Luan, C. Cheng, W. Zhou, and J. Klinowski, *J. Phys. Chem.* 99, 1018 (1995).
48. Z. Luan, H. He, W. Zhou, and J. Klinowski, *J. Chem. Soc., Faraday Trans.* 94, 979 (1998).
49. Z. Luan, M. Hartmann, D. Zhao, W. Zhou, and L. Kevan, *Chem. Mater.* 11, 1621 (1999).
50. Z. Luan, E. M. Maes, P. A. W. van der Heide, D. Zhao, R. S. Czernuszewicz, and L. Kevan, *Chem. Mater.* 11, 3680 (1999).
51. Z. Luan, J. Y. Bae, and L. Kevan, *Chem. Mater.* 12, 3202 (2000).
52. Q. Huo, D. Margolese, and G. D. Stucky, *Chem. Mater.* 8, 1147 (1996).
53. D. Zhao and D. Goldfarb, *Stud. Surf. Sci. Catal.* 97, 181 (1995).
54. R. Ryoo, S. H. Joo, and J. M. Kim, *J. Phys. Chem. B*, 103, 7435 (1999).
55. Dai L. Wang, T. Bu, and L. Chen G., *Colloids Surf. A* 181, 151 (2001).
56. F. Chen, L. Huang, and Q. Li, *Chem. Mater.* 9, 2685 (1997).
57. D. Zhao and D. Goldfarb, *J. Chem. Soc. Chem. Commun.* 875 (1995).
58. A. Corma, Q. B. Kan, and F. Rey, *Chem. Commun.* 5, 579 (1998).
59. J. Xu, Z. Luan, H. He, W. Zhou, and L. Kevan, *Chem. Mater.* 10, 3690 (1998).
60. J. S. Lee, S. H. Joo, and R. Ryoo, *J. Am. Chem. Soc.* 124, 1156 (2002).
61. L. Sierra and J. L. Guth, *Microporous Mesoporous Mater.* 27, 243 (1999).
62. D. Zhao, J. Feng, Q. Huo, N. Melosh, G. H. Fredrickson, B. F. Chmelka, and G. D. Stucky, *Science* 279, 548 (1998).
63. D. Zhao, Q. Huo, J. Feng, B. F. Chomelka, and G. D. Stucky, *J. Am. Chem. Soc.* 120, 6024 (1998).
64. P. Schmidt-Winkel, W. W. Lukens, D. Zhao, P. Yang, B. F. Chmelka, and G. D. Stucky, *J. Am. Chem. Soc.* 121, 254 (1999).
65. B. Tian, X. Liu, Z. Zhang, B. Tu, and D. Zhao, *J. Solid State Chem.* 167, 324 (2002).
66. C. Yu, Y. Yu, and D. Zhao, *Chem. Commun.* 575 (2000).
67. D. Zhao, P. Yang, B. F. Chmelka, and G. D. Stucky, *Chem. Mater.* 11, 1174 (1999).
68. C. Yu, B. Tian, J. Fan, G. D. Stucky, and D. Zhao, *J. Am. Chem. Soc.* 124, 4556 (2002).
69. M. Kruk, M. Jaroniec, C. H. Ko, and R. Ryoo, *Chem. Mater.* 12, 1961 (2000).
70. R. Ryoo, C. H. Ko, M. Kruk, V. Antochshuk, and M. Jaroniec, *J. Phys. Chem. B* 104 11465 (2000).
71. P. I. Ravikovitch and A. V. Neimark, *J. Phys. Chem. B* 105, 6817 (2001).
72. J. Fan, C. Yu, L. Wang, B. Tu, D. Zhao, Y. Sakamoto, and O. Terasaki, *J. Am. Chem. Soc.* 123, 12113 (2001).
73. J. Arupatrakorn and J. D. Tilley, *J. Am. Chem. Soc.* 124, 8380 (2002).
74. K. K. Zhu, Z. N. Ma, Y. Zou, W. Z. Zhou, T. Chen, and H. Y. He, *Chem. Commun.* 24, 2552 (2001).
75. R. Ryoo, S. H. Joo, and S. J. Jun, *J. Phys. Chem. B* 103, 7743 (1999).
76. S. Jun, S. H. Joo, R. Ryoo, M. Kruk, M. Jaroniec, Z. Liu, T. Ohsuna, and O. Terasaki, *J. Am. Chem. Soc.* 122, 10712 (2000).
77. S. H. Joo, S. J. Choi, I. Oh, J. Kwak, Z. Liu, O. Terasaki, and R. Ryoo, *Nature* 412, 169 (2001).
78. S. Joo, S. Jun, and R. Ryoo, *Microporous Mesoporous Mater.* 44, 153 (2001).
79. R. Ryoo, S. H. Joo, M. Kruk, and M. Jaroniec, *Adv. Mater.* 13, 677 (2001).
80. H. J. Shin, R. Ryoo, M. Kruk, and M. Jaroniec, *Chem. Commun.* 4, 349 (2001).
81. M. Kruk, M. Jaroniec, R. Ryoo, and S. H. Joo, *J. Phys. Chem. B* 104, 7960 (2000).
82. M. Kaneda, T. Tsubakiyama, A. Carlsson, Y. Sakamoto, T. Ohsuna, O. Terasaki, S. H. Joo, and R. Ryoo, *J. Phys. Chem. B* 106, 1256 (2002).
83. Z. Liu, Y. Sakamoto, T. Ohsuna, K. Hiraga, O. Terasaki, C. Ko, H. Shin, and R. Ryoo, *Angew. Chem. Int. Ed.* 39, 3107 (2000).
84. M. Huang, A. Choudrey, and P. Yang, *Chem. Commun.* 12, 1063 (2000).
85. W. Zhang, J. Shi, H. Chen, Z. Hua, and D. Yan, *Chem. Mater.* 13, 648 (2001).
86. P. Yang, D. Zhao, D. I. Margolese, B. F. Chmelka, and G. D. Stucky, *Nature* 396, 152 (1998).
87. C. G. Göltner and M. Antonietti, *Adv. Mater.* 9, 431 (1997).
88. B. Tian, H. Yang, X. Liu, S. Xie, C. Yu, J. Fan, B. Tu, and D. Y. Zhao, *Chem. Commun.* 17, 1824 (2002).
89. X. Liu, B. Tian, C. Yu, F. Gao, S. Xie, B. Tu, R. Che, L. Peng, and D. Zhao, *Angew. Chem. Int. Ed.* 41, 3876 (2002).
90. P. Yang, D. Zhao, D. I. Margolese, B. F. Chmelka, and G. D. Stucky, *Chem. Mater.* 111, 2813 (1999).
91. Y. Lu, R. Ganguli, C. A. Drewien, M. T. Anderson, J. C. Brinker, W. Gong, Y. Guo, H. Soyez, B. Dunn, M. H. Huang, and J. I. Zink, *Nature* 389, 364 (1997).
92. D. Zhao, P. Yang, N. Melosh, J. Feng, B. F. Chmelka, and G. D. Stucky, *Adv. Mater.* 10, 1380 (1998).
93. D. Zhao, P. Yang, D. I. Margolese, B. F. Chmelka, and G. D. Stucky, *Chem. Commun.* 22, 2499 (1998).
94. H. Yang, A. Kuperman, N. Coombs, S. Mamiche-Afara, and G. A. Ozin, *Nature* 379, 703 (1996).
95. H. Yang, N. Coombs, I. Sakolov, and G. A. Ozin, *Nature* 381, 589 (1996).
96. S. H. Tolbert, T. E. Schäffer, J. Feng, P. K. Hansma, and G. D. Stucky, *Chem. Mater.* 9, 1962 (1997).
97. Y. F. Lu, H. Y. Fan, N. Doke, D. A. Loy, R. A. Assink, D. A. LaVan, and C. J. Brinker, *J. Am. Chem. Soc.* 122, 5258 (2000).
98. P. Feng, X. Bu, G. D. Stucky, and D. J. Pine, *J. Am. Chem. Soc.* 122, 994 (2000).
99. N. A. Melosh, P. Lipic, F. S. Bates, F. Wudl, G. D. Stucky, G. H. Fredrickson, and B. F. Chmelka, *Macromolecules* 32, 4332 (1999).
100. N. A. Melosh, P. Davidson, and B. F. Chmelka, *J. Am. Chem. Soc.* 122, 823 (2000).
101. H. Yang, Q. Shi, X. Liu, S. Xie, D. Jiang, F. Zhang, C. Yu, B. Tu, and D. Zhao, *Chem. Commun.*, in press.
102. Y. F. Lu, H. Y. Fan, A. Stump, T. L. Ward, T. Rieker, and C. J. Brinker, *Nature* 398, 223 (1999).
103. D. Zhao, Q. Huo, J. Feng, J. M. Kim, Y. J. Han, and G. D. Stucky, *Chem. Mater.* 11, 2668 (1999).
104. J. Fan, C. Yu, and D. Zhao, *Stud. Surf. Sci. Catal.* 135, 286 (2001).
105. J. Sun, J. A. Moulijn, K. C. Jansen, T. Maschmeyer, and M. O. Coppens, *Adv. Mater.* 13, 327 (2001).
106. R. Mokaya, *Angew. Chem. Int. Ed.* 38, 2930 (1999).
107. W. M. Meier, D. H. Olson, and Ch. Baerlocher, Atlas of Zeolite Structure Type, 4th ed.
108. J. M. Kim, J. H. Kwak, S. Jun, and R. Ryoo, *J. Phys. Chem.* 99, 16742 (1995).
109. Z. Zhang, Y. Han, L. Zhu, R. Wang, Y. Yu, S. Qiu, D. Zhao, and F. S. Xiao, *Angew. Chem. Int. Ed.* 41, 2226 (2002).
110. F. Xiao, Y. Han, Y. Yu, X. Meng, M. Yang, and S. Wu, *J. Am. Chem. Soc.* 124, 888 (2002).
111. Z. R. Zhang, R. W. Hicks, T. R. Pauly, and T. J. Pinnavaia, *J. Am. Chem. Soc.* 124, 1592 (2002).
112. W. Guo, L. Huang, P. Deng, Z. Xue, and Q. Li, *Microporous Mesoporous Mater.* 44, 427 (2001); D. Zhao, C. Nie, Y. Zhou, S. Xia, L. Huang, and Q. Li, *Catal. Today* 68, 11 (2001).
113. C. Yu, B. Tian, J. Fan, G. D. Stucky, and D. Zhao, *Chem. Commun.* 2726 (2001).
114. L. Huang, W. Guo, P. Deng, Z. Xue, and Q. Li, *J. Phys. Chem. B* 104, 2817 (2000); S. Che, Y. Sakamoto, O. Terasaki, and T. Tatsumi, *Chem. Mater.* 13, 2237 (2001).

115. R. E. Morris and S. J. Weigel, *Chem. Soc. Rev.* 26, 309 (1997).
116. Q. Huo, J. Feng, F. Schüth, and G. D. Stucky, *Chem. Mater.* 9, 14 (1997).
117. S. Schacht, Q. Huo, G. Voigt-Martin, G. D. Stucky, and F. Schüth, *Science* 273, 768 (1996).
118. Q. Luo, L. Li, Z. Xue, and D. Zhao, *Stud. Surf. Sci. Catal.* 129, 37 (2000).
119. C. J. Brinker, Y. Lu, A. Sellinger, and H. Fan, *Adv. Mater.* 11, 579 (1999).
120. C. Yu, B. Tian, J. Fan, G. D. Stucky, and D. Zhao, *Chem. Lett.* 62 (2002).
121. Y. Yue and Z. Gao, *Chem. Commun.* 18, 1755 (2000); D. M. Antonelli and J. Y. Ying, *Angew. Chem. Int. Ed.* 34, 2014 (1995).
122. U. Ciesla, D. Demuth, R. Leon, P. Petroff, G. D. Stucky, K. Unger, and F. Schuth, *Chem. Commun.* 11, 1387 (1994).
123. F. Vaudry, S. Khodabandeh, and M. E. Davis, *Chem. Mater.* 8, 1451 (1996).
124. P. Feng, Y. Xia, J. Feng, X. Bu, and G. D. Stucky, *Chem. Commun.* 949 (1997); D. Zhao, Z. Luan, and L. Kevan, *Chem. Commun.* 11, 1009 (1997).
125. A. Bhaumik and S. J. Inagaki. *J. Am. Chem. Soc.* 123, 691 (2001).
126. C. Yu, F. Jie, B. Tian, T. Bo, G. D. Stucky, and D. Zhao, *Adv. Mater.*, in press.
127. J. Lee, K. Sohn, and T. Hyeon, *J. Am. Chem. Soc.* 123, 5146 (2001).
128. J. Lee, S. Yoon, T. Hyeon, S. M. Oh, and K. B. Kim, *Commun. Chem.* 21, 2177 (1999).
129. A. Stein, B. J. Melde, and R. C. Schroden, *Adv. Mater.* 12, 1403 (2000).
130. R. C. Schroden, C. F. Blanford, B. J. Melde, B. J .S. Johnson, and A. Stein, *Chem. Mater.* 13, 1074 (2001).
131. B. Lebeau, C. E. Fowler, S. R. Hall, and S. Mann, *J. Mater. Chem.* 9, 2279 (1999).
132. C. E. Fowler, S. L. Burkett, and S. Mann, *Chem. Commun.* 18, 1769 (1997).
133. D. Margolese, J. A. Melero, S. C. Christiansen, B. F. Chmelka, and G. D. Stucky, *Chem. Mater.* 12, 2448 (2000).
134. X. Feng, G. E. Fryxell, L. Q. Wang, A. Y. Kim, J. Liu, and K. M. Kemner, *Science* 276, 923 (1997).
135. K. M. Kemner, X. Feng, J. Liu, G. E. Fryxell, L. Q. Wang, A. Y. Kim, M. Gong, and S. Mattigod, *J. Synchrotron Radiat.* 6, 633 (1999).
136. A. M. Liu, K. Hidajat, S. Kawi, and D. Zhao, *Chem. Commun.* 13, 1145 (2000).
137. Q. H. Yang, M. P. Kapoor, and S. Inagaki, *J. Am. Chem. Soc.* 124, 9694 (2002).
138. S. Inagaki, S. Guan, T. Ohsuna, and O. Terasaki, *Nature* 416, 304 (2002).
139. S. Y. Guan, S. Inagaki, T. Ohsuna, and O. Terasaki, *Microporous Mesoporous Mater.* 44, 165 (2001).
140. A. Fukuoka, Y. Sakamoto, S. Guan, S. Inagaki, N. Sugimoto, Y. Fukushima, K. Hirahara, S. Iijima, and M. Ichikawa, *J. Am. Chem. Soc.* 123, 3373 (2001).
141. T. Asefa, M. Kruk, M. J. MacLachlan, N. Coombs, H. Grondey, M. Jaroniec, and G. A. Ozin, *J. Am. Chem. Soc.* 123, 8520 (2001).
142. T. Asefa, M. J. MacLachan, N. Coombs, and G. A. Ozin, *Nature* 402, 867 (1999).
143. A. Corma, Q. Kan, and M. T. Navarro, *Chem. Mater.* 2123 (1997).
144. A. Sayari, Y. Yang, and M. Kruk, *J. Phys. Chem. B* 103, 3651 (1999).
145. Y. H. Sakamoto, M. Kaneda, O. Terasaki, D. Zhao, J. M. Kim, G. Stucky, H. J. Shim, and R. Ryoo, *Nature* 408, 449 (2000).
146. Y. Sakamoto, I. Diaz, O. Terasaki, D. Zhao, J. Perez-Pariente, J. M. Kim, and G. D. Stucky, *J. Phys. Chem. B* 106, 3118 (2002); J. Miao, T. Ohsuna, O. Terasaki, K. O. Hodgson, and M. A. O'Keefe, *Phys. Rev. Lett.* 89, art. no. 155502 (2002).
147. J. Zhang, Z. Luz, H. Zimmermann, and D. Goldfarb, *J. Phys. Chem. B* 104, 279 (2000).
148. J. Zhang, Z. Luz, and D. Goldfarb, *J. Phys. Chem. B* 101, 7087 (1997).
149. W. W. Lukens, P. Schmidt-Winkel, D. Zhao, J. Feng, and G. D. Stucky, *Langmuir* 15, 5403 (1999).
150. F. Marlow and F. Kleitz, *Micorporous Mesoporous Mater.* 44, 671 (2001); F. Marlow, D. Zhao, and G. D. Stucky, *Micorporous Mesoporous Mater.* 39, 37 (2000).
151. Q. Huo, D. Zhao, J. Feng, K. Weston, S. K. Buratto, G. D. Stucky, S. Schacht, and F. Schüth, *Adv. Mater.* 9, 974 (1997).
152. D. Zhao, J. Sun, Q. Li, and G. D. Stucky, *Chem. Mater.* 12, 275 (2000).
153. P. Yang, D. Zhao, B. F. Chmelka, and G. D. Stucky, *Chem. Mater.* 10, 2033 (1998).
154. I. A. Aksay, M. Trau, S. Manne, I. Honma, N. Yao, L. Zhou, P. Fenter, P. M. Eisenberger, and S. M. Gruner. *Science* 273, 892 (1996).
155. G. Buchel, K. K. Unger, A. Matsumotoand, and K. Tsutsumi, *Adv. Mater.* 10, 1036 (1998).
156. S. M. Yang, N. Coombs, and G. A. Ozin. *Adv. Mater.* 12, 1940 (2000).
157. G. V. R. Rao, G. P. Lopez, J. Bravo, H. Pham, A. K. Datye, H. F. Xu, and T. L. Ward, *Adv. Mater.* 14, 1301 (2002).
158. C. Yu, J. Fan, and D. Zhao. *Acta Chim. Sinica* 60, 1357 (2002).
159. H. Yang, N. Coombs, and G. A. Ozin, *Nature* 386, 692 (1997).
160. J. M. Kim, S. K. Kim, and R. Ryoo, *Chem. Commun.* 2590 (1998).
161. S. Guan, S. Inagaki, T. Ohsuna, and O. Terasaki, *J. Am. Chem. Soc.* 122, 5660 (2000).
162. H. Lin and C. Y. Mou, *Science* 273, 756 (1996).
163. H. Lin, C. Y. Mou, and S. Liu, *Adv. Mater.* 12, 103 (2000).
164. S. S. Kim, W. Zhang, and T. J. Pinnavaia, *Science* 282, 1302 (1998).
165. V. Y. Gusev, X. B. Feng, Z. Bu, G. L. Haller, and J. A. O'Brien, *J. Phys. Chem.* 100, 1985 (1996).
166. T. Tatsumi, K. A. Koyano, Y. Tanaka, and S. Nakata, *J. Porous Mater.* 6, 13 (1999).
167. M. Hartmann and C. Bischof, *J. Phys. Chem. B* 103, 6230 (1999).
168. M. A. Springuel-Huet, J. L. Bonardet, A. Gedeon, Y. Yue, V. N. Romannikov, and J. Fraissard, *Microporous Mesoporous Mater.* 44, 775 (2001).
169. K. Cassiers, T. Linssen, M. Mathieu, M. Benjelloun, K. Schrijnemakers, P. Van Der Voort, P. Cool, and E. F. Vansant, *Chem. Mater.* 14, 2317 (2002).
170. M. Broyer, S. Valange, J. P. Bellat, O. Bertrand, G. Weber, and Z. Gabelica, *Langmuir* 18, 5083 (2002).
171. M. Hartmann and A. Vinu, *Langmuir* 18, 8010 (2002).
172. H. Van Bekkum and A. J. Hoefnagel, *Stud. Surf. Sci. Catal.* 83, 397 (1994).
173. K. P. Kloetstra and H. Van Bekkum, *J. Chem. Res.* 26, 234 (1995).
174. K. P. Kloetstra and H. Van Bekkum, *J. Chem. Soc., Chem. Commun.* 1005 (1995).
175. I. V. Kozhevnikov and A. Sinnema, *Catal. Lett.* 30, 241 (1995).
176. A. Corma and A. Martinez, *J. Catal.* 153, 25 (1995).
177. Y. S. Ko, T. K. Han, J. W. Park, and S. I. Woo, *Macromol. Rapid Commun.* 379, 703 (1996).
178. J. Tudor and D. O'Hare, *Chem. Commun.* 603 (1997).
179. K. Kageyama, J. Tamazawa, and T. Aida, *Science* 285, 2113 (1999).
180. F. Marlow, M. D. McGehee, D. Zhao, B. F. Chmelka, and G. D. Stucky, *Adv. Mater.* 11, 632 (1999).
181. P. D. Yang, G. Wirnsberger, H. C. Huang, S. R. Cordero, M. D. McGehee, B. Scott, T. Deng, G. M. Whitesides, B. F. Chmelka, S. K. Buratto, and G. D. Stucky, *Science* 287, 465 (2000).
182. G. Wirnsberger and G. D. Stucky, *Chem. Mater.* 12, 2525 (2000).
183. O. Dag, G. A. Ozin, and H. Yang, *Adv. Mater.* 11, 474 (1999).
184. L. F. Karen, M. H. Bartl, A. Popitsch, and G. D. Stucky, *Angew. Chem. Int. Ed.* 41, 959 (2002).
185. M. H. Huang, A. Choudrey, and P. Yang, *Chem. Commun.* 1063 (2000).
186. Y. Han, J. M. Kim, and G. D. Stucky, *Chem. Mater.* 12, 2068 (2000).

187. Z. Liu, Y. Sakamoto, T. Ohsuna, O. Terasaki, C. H. Ko, H. J. Shin, and R. Ryoo, *Angew. Chem. Int. Ed.* 39, 3107 (2000).
188. C. H. Ko and R. Ryoo, *Chem. Commun.* 2467 (1996).
189. S. Inagaki, Y. Fukushima, and K. Kuroda, *J. Chem. Soc., Chem. Commun.* 680 (1993).
190. S. Inagaki, S. Guan, Y. Fukushima, T. Ohsuna, and O. Terasaki, *J. Am. Chem. Soc.* 121, 9611 (1999).
191. M. Sasaki, M. Osada, N. Sugimoto, S. Inagaki, Y. Fukushima, A. Fukuoka, and M. Ichikawa, *Microporous Mesoporous Mater.* 21, 597 (1998).
192. A. Fukuoka, Y. Sakamoto, S. Guan, S. Inagaki, N. Sugimoto, Y. Fukushima, K. Hirahara, S. Iijima, and M. Ichikawa, *J. Am. Chem. Soc.* 121, 9611 (2001).
193. J. Liu, X. Feng, G. E. Fryxell, L. Wang, A. Y. Kim, and M. L. Gong, *Adv. Mater.* 10, 161 (1998).
194. J. Brown, L. Mercier, and T. J. Pinnavaia, *Chem. Commun.* 69 (1999).
195. S. Dai, M. C. Burleigh, Y. H. Ju, H. J. Gao, J. S. Lin, S. J. Pennycook, C. E. Barnes, and Z. L. Xue, *J. Am. Chem. Soc.* 122, 992 (2000).
196. Y. J. Han, G. D. Stucky, and A. Butter, *J. Am. Chem. Soc.* 121, 989 (1999).
197. E. Yablonovitch, *Phys. Rev. Lett.* 58, 2059 (1987).
198. O. D. Velev, T. A. Jede, R. F. Lobo, and A. M. Lenhoff, *Nature* 389, 447 (1997).
199. Q. Luo, Z. Liu, L. Li, S. Xie, J. Kong, and D. Zhao, *Adv. Mater.* 13, 286 (2001).
200. J. E. G. J. Wijnhoven and W. L. Vos, *Science* 281, 802 (1998).
201. B. T. Holland, C. F. Blanford, and A. Stein, *Science* 281, 538 (1998).
202. J. Wijinhoven and W. L. Vos, *Science* 281, 802 (1998).
203. A. A. Zakhidov, R. H. Baughman, Z. Iqbal, C. Cui, I. Khayrullin, S. O. Dantas, J. Marti, and V. G. Ralchenko, *Science* 282, 897 (1998).
204. A. Blanco, E. Chomski, S. Grabtchak, M. Ibisate, S. John, S. W. Leonard, C. Lopez, F. Meseguer, H. Miguez, J. P. Mondia, G. A. Ozin, O. Toader, and H. M. V. Driel, *Nature* 405, 437 (2000).
205. A. Imhof and D. J. Pine, *Adv. Mater.* 10, 697 (1998).
206. P. V. Braun and P. Wiltzius, *Nature* 402, 603 (1999).
207. S. A. Davis, S. L. Burkett, N. H. Mendelson, and S. Mann, *Nature* 358, 420 (1997).
208. A. Imhof and D. J. Pine, *Nature* 389, 948 (1997).
209. S. Shen, Y. Li, Z. Zhang, J. Fan, B. Tu, W. Zhou, and D. Zhao, *Chem. Commun.* 19, 2212 (2002).
210. P. T. Tanev and T. J. Pinnavaia, *Science* 271, 1267 (1996).
211. P. T. Tanev and T. J. Pinnavaia, *J. Am. Chem. Soc.* 119, 8616 (1997).
212. S. S. Kim and T. J. Pinnavaia, *Chem. Commun.* 23, 2418 (2001).
213. A. Navrotsky, I. Petrovic, Y. Hu, C. Y. Chen, and M. E. Davis, *Microporous Mater.* 4, 95 (1995).
214. Mobil Company, Communication in 3rd IMMS, Jeju, Korea, 2002.

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanopowders Produced Using Microreactors

Rainer Schenk, Volker Hessel

Institut für Mikrotechnik Mainz GmbH, Mainz, Germany

Nathalie Jongen

Swiss Federal Institute of Technology, Lausanne, Switzerland

Vincenzo Buscaglia

Institute of Energetics and Interfaces—National Research Council, Genova, Italy

Sophie Guillemet-Fritsch

Université Paul Sabatier, Toulouse, France

Alan G. Jones

University College London, London, United Kingdom

CONTENTS

1. INTRODUCTION

One of the most widely used classes of inorganic powders is ceramic powders. Some conventional applications such as pigments, catalysts, abrasives, polishing agents, and lubricants contain ceramic powders. An additional area of application includes provision of precious materials for electrical and electronic devices such as capacitors, piezoelectric transducers, and thermistors just to name a few [1]. Furthermore, ceramic powders are also applied as mechanical, structural material for engine components, cutting tools, and catalytic converters. Additionally, coatings to enhance corrosion resistance, thermal barrier, and electrical insulation are well-known areas for the use of ceramic powders.

Beside ceramic powders, there is a need for organic powders in the life-science industry for medical and cosmetic applications. A further large-scale application is the area of pigments and paints for the automotive industry, for household utensils, and in the furniture industry [2]. New developments in this area concern the application of polymer powders in aqueous or solvent-containing dispersions for paint and coating purposes.

One of the most common processes for powder production is the precipitation of the powders in large stirred tanks. This process allows the cost-effective production of powders with a wide variation of chemical composition in huge quantities. In the future there will be an increasing need for nanoparticles with well-defined size, such as for the construction of electronic devices with unique properties

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (287–296)

concerning optical band gap [3]. As these properties are ultimately determined by the physical size and shape of the crystallites, new processing routes for powders are required.

Supersaturation, the driving force for precipitation, is caused by feeding a second reactant, and crystallization frequently starts in an uncontrolled way at the feeding point of one of the reaction partners. In batches, these processes regularly happen because the reaction solution added makes contact with the full concentrated solution of the second reaction partner and reacts immediately [4]. Additionally, industrial crystallizers provide extremely unequal flow conditions, with local velocities, shear rates, and energy dissipation rates varying by orders of magnitude throughout the vessel [5]. Many extremely irregular small crystals are created resulting in the formation of particles with a very wide range of the particle size distribution [4]. Furthermore, the scale-up from a small laboratory vessel to a large production tank is often costly, time-consuming, and a loss of powder quality is observed. In several cases, this scale-up failed [6].

To overcome these drawbacks it is necessary to improve the mixing quality of the reactants to achieve a homogeneous distribution of the compounds in the reactor and to install a continuous process. Hence, the residence time of the particles in the reactor can be controlled and optimized for the application. Additionally, for an enhanced throughput, the number of the continuously working reactors can be increased, and the highly risky scale-up step can be avoided.

A recent concept for powder production is based on the combination of microreaction technology with a tubular reactor. Micromixers are well known for their highly efficient, fast, and homogenous mixing of gases [7] and liquids [8]. Two of the most common concepts are based on multilamination, for example realized in interdigital mixers, or split and recombined of fluid flows [8]. Multilamination is a simple and reliable concept, as it relies only on alternatingly arranged lamellae. Two fluids are divided in several substreams of thicknesses beyond 100 μm in such a way that regular and alternating lamellae of the two different fluids are obtained. For fast mixing, these lamellae can be focused geometrically in a special chamber by a factor of 50 or more so that mixing within a few milliseconds can be achieved [9].

In this context, it should be mentioned that the expression " microreactor" is not well defined, for example comprising microchips, complete table-top units, large reactors of shoe-box size, or smaller ones of match-box size. In this review article, " microreactor" is used as term for a chemical reactor of undefined size that contains components with microstructured channels or tubes in the microscale. In general, the width of such channels is in the micrometer range, mostly between 50 to 500 μm [8]. In the field of powder technology, the term " microreactor" is sometimes used in a completely different way, referring to supramolecular aggregates or so. "Micelles" [e.g., 10, 11] or "spherulites" [12] in microemulsions or miniemulsions [13] as well as salt particles in a modified aerosol decomposition process [14] were also named " microreactors." It is beyond the scope of this review to include the above-mentioned methods for the generation of nanoparticles.

In general, the use of microreactors has the advantage over conventional reactors of fast and homogenous mixing [8, 15] of the reactants as well as efficient heat transfer, which allows fast heating and cooling cycles [16] and isothermal reaction conditions [17]. Additionally, continuous operation of the reactors can be achieved. In recent years, different microcomponents were developed for operations mixing [8, 15] and heat exchange [18] as well as whole reaction systems for chemical synthesis for thermal [17], heterogeneous catalysis [19–21], photochemical [22, 23], and electrochemical [24–26] processing. Additionally, the small internal dimensions of the reaction channels (width below 200 μm) allow one to perform reactions in the explosion regime [27, 28], which conventionally are not possible. For example, the direct fluorination of toluene with elemental fluorine selective to monofluorotoluenes was demonstrated using a falling film microreactor [29]. Hence, microreactors are now well established in the fields of gas, gas/liquid, and liquid phase reactions. In this review, the novel approach using microreaction technology for the production of micro- and nanopowders is described.

2. CONVENTIONAL CONTINUOUS PROCESS

A continuous production process can be achieved conventionally, such as with a continuous working crystallizer or with a reactor based on the concept of a "tubular reactor" [4]. For example, a special type of crystallizers allows the continuous precipitation and reaction crystallization in agitated crystallizers. For hydroxides, especially nickel hydroxides, spherical particles with a medium diameter larger than 10 μm can be produced. SEM photographs and diagrams of particle size distributions are given in [4].

3. STATIC MIXER-TUBULAR REACTOR SETUP

As mentioned in the Introduction, a first step to overcoming the drawbacks of stirred tank reactors is the use of a continuous operating tubular reactor equipped with a static mixer [30]. The static mixer contains bow-tie shaped helical static mixing elements with clockwise and counterclockwise orientations with an internal width of 14.3 mm. $BaTiO_3$ powder with a mean particle size of 80 nm and a specific surface area of 25 m^2/g was synthesized using this static mixer-tubular reactor setup [31]. An X-ray powder diffraction pattern shows that these $BaTiO_3$ particles are single phased. A 3 mol/l NaOH solution as well as a solution containing $BaCl_2$ and $TiCl_4$ with a molar ratio of 1:1.1 were mixed and reacted at 85 °C with an overall flow rate of 360 ml/min, leading to a production of 25 kg $BaTiO_3$ powder per day.

4. MICROREACTOR FOR THERMAL DECOMPOSITION OF A PRECURSOR

A nanoparticle microreactor [32] is a small-scale system capable of converting an aerosol precursor to solid particles. A liquid precursor is evaporated and transported with a carrier gas to a tube-type reactor (diameter = 2 mm, length = 400 mm). In this tube, the precursor is decomposed by heat

and the resulting particles are collected at the outlet of the reactor on a transmission electron microscopy (TEM) grid.

Titanium dioxide particles were synthesized by thermal decomposition of titanium tetraisopropoxide in a nitrogen carrier gas. 1 μl of the precursor was evaporated and subsequently decomposed in the reactor. After a run time of one minute, very small primary titanium dioxide particles with a mean size of 9.0, 20.6, and 31.7 nm could be obtained at a reactor temperature of 500, 700, and 1000 °C, respectively. The small size of the primary particles was explained with a dominating surface reaction at the wall. Increasing the tube diameter to 4 mm leads to larger particles, and the particle size decreases with increasing temperature. The latter result is comparable with the TiO_2 synthesis in a larger tubular flow reactor (tube diameter = 22 mm) [33].

5. MICROMIXER-TUBULAR REACTOR SETUP

The synthesis of cadmium sulfide nanoparticles in a micromixer-tubular reactor setup was reported recently [34]. In this setup, an interdigital micromixer [35] is used where the fluid streams of the two reactants are divided in 16 substreams each. Aqueous solutions of cadmium nitrate and sodium sulfide were delivered to the microfluidic channels at various flow rates (10–300 μl/min). The outlet flow was connected to a quartz flow cell, and the absorption spectra were obtained continuously. All judgments concerning the powder characteristics result from these absorption spectra: Miniaturizing the reaction volume is sufficient to lower the polydispersity of the crystallites, and increasing the volumetric flow rate leads to a further reduction of polydispersity.

6. SEGMENTED FLOW TUBULAR REACTOR (SFTR)

An innovative approach for the production of homogenous fine powders is the continuous operation in a plug-flow reactor, the segmented flow tubular reactor (SFTR) (see Fig. 1) [36]. This apparatus is mainly built from the components mixer, segmenter, tubular section, and separator. The two reactant solutions are mixed in the mixer, forming a supersaturated solution of the powder compound. In the segmenter, the reactant mixture flow is segmented by a nonmiscible fluid so that small volumes of about 100 μl are formed and guided through the tubular section. In the tubular section, the powder is precipitated in the small batch reactors. When those reach the end of the tubular section, the powder suspension is separated from the nonmiscible fluid, which is then recycled. Due to the defined mixing in the mixer and the defined residence time of the reaction solution in the tubular section, powders with narrow particle distribution are obtained. Examples of powders produced within the SFTR are given in the following sections.

In a reactor, the mixing conditions of the two reactants severely influence the resulting quality of the precipitated powder, especially the crystal morphology, the mean particle size, and the particle size distribution [e.g., 37, 38].

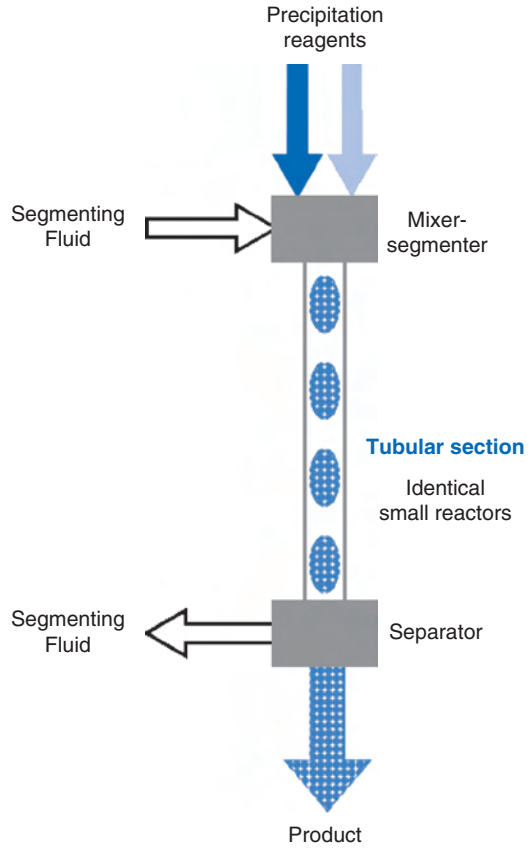


Figure 1. Scheme of the segmented flow tubular reactor (SFTR).

6.1. Hydrodynamics and Mixing Characterization of Micromixers

Microstructured mixers specially suited for the avoidance of fouling were evaluated as offering new means for chemical processes to benefit from micromixing physics. In the case of a residence time of the supersaturated solution in the mixer longer than the nucleation time of the powder, unwanted precipitation of the powder, so-called fouling, can occur in the mixer. For this purpose, new special-type micromixers, designed with regard to antifouling processing, were developed based on following concepts:

- Mixing in the segmenter outside the mixer ("chamber-free mixing") using an impinging-jet micromixer [39].
- Separation of the two reaction solutions by pure water ("delay-type mixing") [40, 41]. (Two different kinds of separation-layer mixers, "concentric" and "stacked," were developed, and their performance was investigated with the precipitation of calcium carbonate.)

Dependent on the flow rate, different types of flow patterns are found for the impinging jet mixers [39]. At low flow rates, the two fluids merge immediately after leaving the outlets, wet the mixer surface, and finally form droplets. At increasing flow rate, the frequency of the droplets increases and results in merging to a jet. The obtained "Y-type" stable jet at a standstill in the normal position is the preferred process regime (Fig. 2a). Increasing flow rate further yields

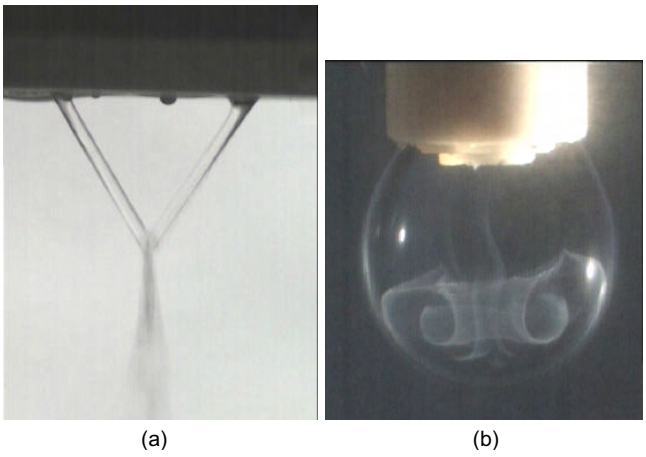


Figure 2. Hydrodynamics of (a) the impinging-jet and (b) concentric separation-layer micromixer.

a broadening of the jet, first termed "fan-shaped," thereafter "fanned-out" when broadening is intense. The stable "Y-type" process regime can be adapted to different flow rates by changing the width of the outlet hole. The mixing quality analyzed with competitive redox reactions, first applied by Villermaux et al. [42], increases with decreasing size of the outlet hole. The best results were obtained with the smallest openings of 350 μm.

Both droplet and jet operation are feasible with the concentric separation-layer micromixer. Two laminar flowing liquid shells surround a liquid core. This situation, however, is perturbed by a more intense mixing process when the three liquids enter the droplet end cap. Here, precipitation occurs (see Fig. 2b). By setting concentration and liquid flow rates, it is possible to avoid contact of the mixed flow with the mixer surface again. However, this generally seems to be more subtle as compared to impinging-jet operation. The mixing quality of the concentric separation-layer mixer is rather poor compared to the impinging-jet and stacked separation-layer micromixer.

In the case of the stacked separation-layer micromixer, the hydrodynamics in aqueous drops formed at the outlet of the rectangular mixing chamber were visualized by the formation of iron rhodanide. The photograph in Figure 3 shows the outlet of the mixer (top in picture) with a drop growing at the outlet and a second drop falling down. Contact of the two reactant solutions leads to the immediate formation of the red iron rhodanide immediately. Accordingly, the finding of a colorless solution inside the drop proves that no contact between the two aqueous reactant solutions (0.05 molar $Fe(NO_3)_3$ and 0.05 molar NaSCN) has taken place. The streamlines in the drop at the top in Figure 3 showed that the mixing has started. The homogeneous gray tone (in reality: red color) inside the second drop (Fig. 3) indicates that 0.4 s later the mixing is finished. By setting suitable flow rates (reactant solutions: 30 ml/h each; water: 200 ml/h), wall contact of the iron rhodanide can be avoided. The mixing quality of the stacked separation layer micromixer is good and increases with increasing water flow rate by a constant flow rate of the two reactants of 250 ml/h each. In the case of an overall flow rate of 1000 ml/h, the mixing quality is comparable with other micromixers such as the interdigital and caterpillar micromixer.

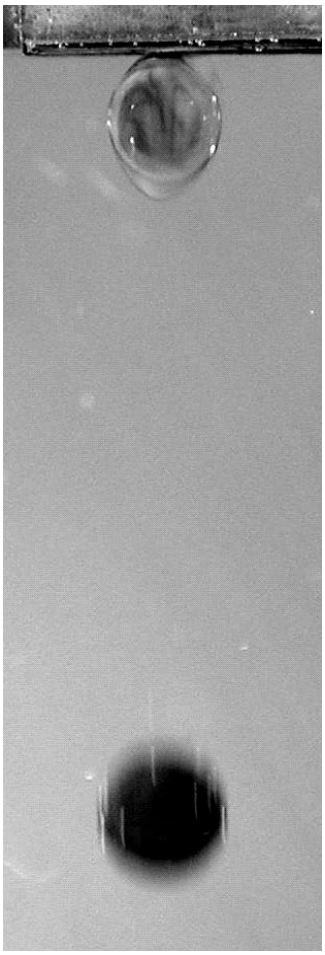


Figure 3. Photograph documenting the hydrodynamics in aqueous drops formed at the outlet of the stacked separation-layer mixer.

6.2. Numbering-Up Concept of the SFTR

For production purposes, the output of the reactor has to be increased. In highly nonlinear systems such as is given in precipitation reactors, scale-up is better ensured when all critical parts of the process are exactly the same at all scales [43]. Therefore, the concept of scaling by replication instead of the conventional scaling by dimensional changes was chosen to increase the productivity of the SFTR technique to avoid the drawbacks encountered with the scale-up procedure [44]. By this the productivity in terms of throughput is enhanced without changing the microfluidic flow properties at all. At present, a pilot plant with six parallel reactors is being built up. Each individual tubular reactor of the pilot plant SFTR is run with the some conditions as in the single-channel SFTR discussed above. The powder characteristics achieved with the continuous laboratory scale SFTR are expected to be maintained using the pilot plant SFTR. This pilot plant will be evaluated with the powder systems

calcium carbonate, barium titanate, and nickel manganese oxalate.

6.3. Nano- and Micropowders Produced within the SFTR

6.3.1. Calcium Oxalate: $Ca(C_2O_4)$

The precipitation of calcium oxalate serves as model system for mechanistic studies [45, 46] because of the different morphologies related to the three hydrates. Additionally, the understanding of the precipitation mechanism of calcium oxalate is important concerning the aim of prevention of the formation of renal calculi [47].

The mono-, di-, and trihydrate forms of calcium oxalate precipitate in the form of dendrites, bipyramids, and long platelets, respectively. These different morphologies are related to the three mineralogical phases [48], whereas in renal calculi the calcium oxalate monohydrate is the dominating morphology [49]. The proportion of calcium oxalate dihydrate (COD) increases with increasing supersaturation and decreases with the mixing time. It is proposed that a supersaturation limit for nucleation of the COD phase exists, and at low level of supersaturation, calcium oxalate monohydrate (COM) is formed [46].

In the case of precipitation in the SFTR equipped with a Vortex mixer, an additional type of aggregate could be achieved [47]. Key factor in the precipitation step is the mixing step influencing the morphology and the phase. The SFTR allows the observation of second-level aggregates, which are destroyed in a conventional batch reactor due to the shear forces induced by the impeller. The static flow conditions in the SFTR during the precipitation step result in an undisturbed formation of agglomerates [47].

6.3.2. Yttrium Barium Oxalate: $Y_2Ba_4(C_2O_4)_7$

Mixed yttrium barium oxalates are suitable precursors for the synthesis of the corresponding oxides through high-temperature pyrolysis. These mixed oxides are of interest in the area of high-temperature superconducting materials.

In a batch reactor under all investigated conditions, there is always a mixture of $Y_2Ba_4(C_2O_4)_7 \cdot 18_2O$ and $Y_2Ba_2(C_2O_4)_5 \cdot 8H_2O$ obtained [50]. The oxalate $Y_2Ba_4(C_2O_4)_7 \cdot 18H_2O$ is preferably formed at low temperature (30 °C) with fast addition of reactants, small volume of seed suspension, and a small mixing volume. At higher temperature (40 °C), slow reactant addition, high volume flow of the seed suspension, and a large mixing volume favors the formation of the mixed oxalate $Y_2Ba_2(C_2O_4)_5 \cdot 8H_2O$, which crystallizes in a pinacoidal shape. In contrast to the finding in the batch reactor, a selective synthesis of $Y_2Ba_4(C_2O_4)_7 \cdot 18H_2O$ could be performed using the SFTR. $Y_2Ba_4(C_2O_4)_7 \cdot 18H_2O$ forms octahedral crystals below 1 μm in size under the experimental conditions applied. The Ba/Y ratio is experimentally determined to 2.02, indicating that only this phase is precipitated. This result is in agreement with the X-ray diffractogram, which shows only the peaks that correspond to $Y_2Ba_4(C_2O_4)_7 \cdot 18H_2O$, and the peaks for the compound $Y_2Ba_2(C_2O_4)_5 \cdot 8H_2O$ are absent.

6.3.3. Zinc Sulfide: ZnS

Zinc acetylacetonate (0.05 mol/l) was mixed with thioacetamide (0.4 mol/l), which is thermally decomposed to form sulfide anions. In the tubular section, a reaction temperature of 80 °C and a residence time of 30 min was chosen. Comparable spherical nanoparticles were obtained performing the synthesis in a small-scale laboratory batch reactor and in the SFTR. The mean particle size was determined to be 120 nm with a span of 1.4 as well as the specific surface area (SFTR) of 90 m^2/g, and the powder obtained in a batch process was measured to a mean particle size of 90 nm with a span of 1.6 as well as specific surface area of 95 m^2/g. [51]. These investigations show that the ZnS nanopowders can be produced on a large scale by transferring the reaction from a small-scale laboratory batch reactor to a continuous process.

6.3.4. Yttrium Hydroxy Carbonate: $Y(OH)CO_3$

$Y(NO_3)_3$ (0.025 mol/l) and urea (0.75 mol/l) were mixed. In the tubular reactor the urea was thermally decomposed forming OH^- and CO_3^{2-} ions at a temperature of 95 °C. A residence time of 120 min and flow rate of 36 ml/h resulted in spherical particles of a mean size of 0.69 μm. The particles obtained from a small-scale batch process show nearly the same mean particle size of 0.67 μm [51]. In both cases a similar particle size distribution was obtained. The use of a mixer/tubular reactor setup without segmentation of the aqueous phase results in particles with a significantly larger varying particle size distribution.

6.3.5. Copper Oxalate: $Cu(C_2O_4)$

Copper oxalate was synthesized from $Cu(NO_3)_2$ and $Na_2(C_2O_4)$ aqueous solutions in a small-scale batch reactor as well as in the SFTR [36]. After the initial nucleation, two different growth processes, growth of crystallites and aggregation, were observed. The copper oxalate particles exhibit a rod- or cube-like shape. The single crystallites as well as the aggregates vary in size depending on the experimental conditions. At a high supersaturation level, the particle size distribution is narrower using the SFTR, while at a low supersaturation level the batch process delivers powder with a smaller particle size distribution. For production purposes, a high concentration of the reactants in water is preferable. Hence, running the system at a high supersaturation level is advantageous. In the case of copper oxalate synthesis, the type of the reactor has no influence on the powder morphology, in contrast to the addition of additives like hydroxymethylpropylcellulose, which induces a change of the powder morphology.

A further improvement of the favorable characteristics of the SFTR was expected when micromixers, instead of conventional Vortex or T-type mixers, are used. Hence, the SFTR was equipped with an interdigital micromixer. The resulting copper oxalate particles had a cushion-like shape and a mean particle size of 2.9 ± 1 μm [52].

6.3.6. Calcium Carbonate: $CaCO_3$

High-performance $CaCO_3$ powders are used for applications in paint, textile, plastic, adhesive, rubber, bioceramic, and paper industries [53]. Such industrial applications need

well-defined $CaCO_3$ powders with particles of defined crystallinity and morphology, uniform shape, and a narrow particle size distribution. These product properties are essential, and it is necessary to control such powder characteristics [54]. Furthermore, blocking free processing of $CaCO_3$ is a challenging task due to the low solubility of $CaCO_3$ in water. Hence, this reaction is an ideal candidate for fouling investigations [55].

Aqueous solutions of ammonium carbonate and calcium chloride were injected into the SFTR equipped with conventional "V"- or "T"-type mixers [53]. Alternatively, gaseous carbon dioxide or an aqueous solution of calcium hydroxide can be used as the calcium source. For comparison, in a small-scale laboratory batch reactor a mixture of micrometer-sized spherical and brick-shaped particles were obtained. These morphologies are related to vaterite and calcite crystals, respectively.

The formation of calcite or vaterite phases is strongly influenced by experimental conditions like temperature, pH value, the choice of reactants, reactant concentrations, use of additive, and the type of mixer. The competitive phase precipitation of calcium carbonate is very complex, and the synthesis of pure morphologies is only possible in a narrow process window.

Recently, it was reported [56] that the morphology of calcium carbonate can be selected by the use of seeds. In this case, the crystallographic structure of the seeds determines the structure of the powder: calcite seeds lead to calcite particles, whereas vaterite seeds result in the synthesis of a vaterite powder.

For these investigations calcium nitrate (20 mmol/l) and potassium carbonate (20 mmol/l) were used as reactants, and poly(acrylic acid) was added to stabilize the seeds in the reaction solution. Additionally, the SFTR was equipped with an impinging-jet micromixer. This new equipment and the more sophisticated reaction conditions allowed the synthesis of pure calcite powder of a rice-shaped morphology (see Fig. 4) with particles of a smaller order of magnitude (mean particle size = 0.4 μm) than described above and with a narrow particle size distribution (span = 1.06). The specific surface area was determined as being 35.2 m^2/g. In a 24 h continuous experiment, 25 L of a product suspension with 46 g $CaCO_3$ was produced. The calcium carbonate quality was constant concerning the parameters mean particle size, particle size distribution, specific surface area, and morphology. Furthermore, the synthesis of $CaCO_3$ in a small-scale laboratory batch reactor and in the SFTR results in the same powder properties. This allows the optimization of the reaction conditions in a small-scale batch reactor and, for production purposes, the transfer of these conditions to the continuous running SFTR.

Furthermore, computational fluid dynamics have been applied to investigate the mixing and segmentation conditions. Due to the short induction times (<50 ms), seeding has been applied to ensure a stable precipitation process in the SFTR. In the seeded precipitation of calcium carbonate, particles with a regular morphology and a smaller particle size compared to the unseeded case were obtained. Additionally, the predicted particle sizes show good agreement with the experimentally observed particle sizes [57].

6.3.7. Mixed Manganese Nickel Oxalate: $Mn_\alpha Ni_{1-\alpha}(C_2O_4)$

Ceramics made of mixed manganese nickel oxides have a negative temperature coefficient (NTC) and are used as NTC thermistors. The simple sintering of a mixture of the oxide powders is not sufficient to obtain a homogenous distribution of the metals in the ceramic. Hence, the electrical properties are unsuitable for industrial applications [58]. A suitable route to the nickel manganites is the thermal decomposition of the corresponding oxalates, which were synthesized by coprecipitation.

In an SFTR equipped with a "V"-type mixer, manganese and nickel nitrates as well as ammonium oxalate react to the nickel manganese mixed oxalates. Uniform and regular octahedral crystals with a monomodal particle size distribution were precipitated (see Fig. 5) [59]. The mean particle size was 12.5 μm with a span of 0.9. A continuous powder production with a rate of 3 g powder per hour over a period of more then 20 hours was achieved without change of the

Figure 4. SEM picture of a calcium carbonate powder with the SFTR.

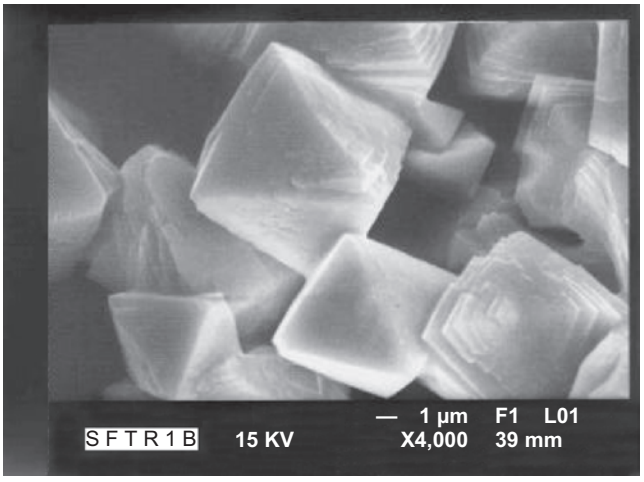


Figure 5. SEM picture of a mixed nickel manganese oxalate powder produced in the SFTR.

morphology of the particles, mean particle size, or size distribution. The corresponding synthesis in a small-scale laboratory batch reactor delivers particles with an irregular shape. The same mean particle size (11.9 μm) was observed, but the particle size distribution (span = 1.5) was significantly larger compared to the particles produced with the SFTR technique.

The above described SFTR process can also be used to produce a mixed nickel manganese cobalt oxalate of the composition $Mn_{0.5}Ni_{0.24}Co_{0.26}(C_2O_4)$. Homogenous and regular octahedral crystals with a mean particle size of 10.8 μm and a span of 0.84 were obtained. For comparison, the particles produced in a batch pilot reactor are irregularly shaped due to inhomogeneous crystal growth zones in the conventional batch reactor [58]. Consequently, the particle size distribution is enlarged (span = 1.31), whereas the mean particle size (10.4 μm) is comparable to the powder produced with the SFTR process.

6.3.8. Barium Titanate: $BaTiO_3$

Barium titanate is widely used in ceramic capacitors [60] as dielectric material in multilayer chip capacitors (MLC) due to the fact that the dielectric constant can be manipulated to increase capacitance. The higher dielectric constant ceramics usually have higher temperature and voltage sensitivity of capacitance. The future trend of MLC development goes in the direction of reduced layer thickness and increased number of layers; at present, layer thicknesses of 10 to 20 μm are common but will be decreased to approximately 2 μm. Recently, the area of application of ceramic capacitors is enlarged to switched mode power supplies (SMPS). Ceramic capacitors allow the design of smaller sized components for smaller package sizes in the SMPS itself [60]. An additional application for barium titanate is as a semiconductor material in PCTR (positive temperature coefficient of resistivity) devices [61, 62].

Barium titanate was synthesized in an SFTR equipped with a conventional "V" mixer or with a caterpillar micromixer. A solution containing $BaCl_2$ and $TiCl_4$ was mixed with a NaOH solution at room temperature. Subsequently, the precipitation of barium titanate occurs in the tubular section at a temperature of 80 °C to 95 °C. A residence time of about 20 min was achieved [63].

The mean particle size of barium titanate decreases with increasing concentration of the chloride solution. At the maximum reactant concentration, spherical particles of approximately 30 nm in size (SEM observation) were obtained [64]. This value is in good agreement with the surface area of 40 m^2/g and the crystallite size of 30 nm measured with X-ray diffraction. Figure 6 shows particles with a mean size of about 100 nm produced at a concentration of the reaction solution of 0.2 mol/l. At a chloride concentration of 0.062 mol/l, the mean particle size was determined as being 400 nm using a "V"-mixer and to 200 nm using a caterpillar micromixer, respectively.

For comparison, it should be mentioned that powders produced by hydrothermal methods can meet the requirements concerning the production of nanosized particles (30–60 nm), but this fine powder produced in a batch reactor suffers from an insufficient stoichiometric control and intensive agglomeration. The graph of the particle size distribution of fine commercial powders is given in [63]. Both commercial powders are composed of primary particles with a size of about 60 nm. The grade of agglomeration differs quite a bit so that 55% of the particles of one of the commercial powders is smaller than 100 nm, while only 25% of the other commercial powder particles are below a size of 100 nm.

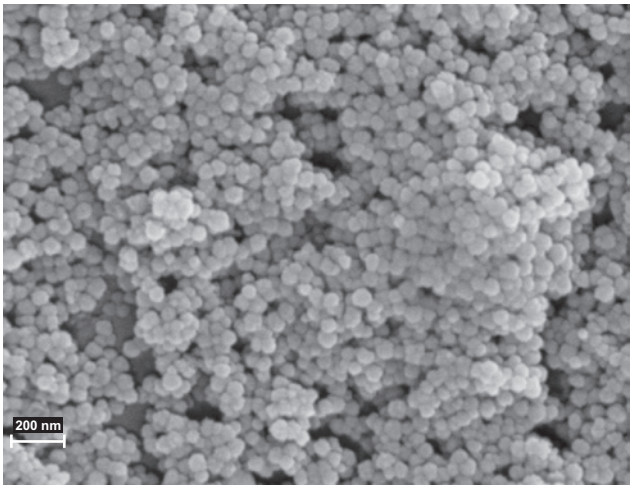


Figure 6. SEM picture of barium titanate powder produced with the SFTR.

The degree of agglomeration of $BaTiO_3$ powders was determined for the powder produced using the SFTR and was compared with powder produced according to the low temperature aqueous synthesis (LTAS) method in a 6-liter batch reactor [62]. Both reactors produce nanometer sized primary particles with a medium diameter of 23 nm for the SFTR and 30 nm for the LTAS system. Subsequent freeze drying and redispersion of these powders lead to agglomeration, but the SFTR powder is of superior quality compared to LTAS powder: 80% (57%) of the particles are smaller than 100 nm if produced with the SFTR (LTAS), respectively.

7. MICROJET REACTOR

The principle of this reactor is relatively simple: two reactant solutions were injected with high velocity (up to 1,000 m/s) into a reaction chamber through a small nozzle (between 100 and 1,000 μm), and colloids are formed in the surrounding gas [65, 66]. Fine spray is generated, which results in a fast and homogenous mixing in the fine droplets. The powder precipitates in these fine droplets immediately. Additionally, a gas stream is injected into the reaction chamber, which allows cooling of exothermic reactions. Furthermore, this gas stream removes the precipitated powder out of the reaction chamber completely.

One application of the MicroJet Reactor is the precipitation of organic pigments [67]. For example, the pigment tetrachloro copper phthalocyanine was precipitated in the MicroJet reactor. The mean particle size was determined as being 26 nm with a standard deviation of 11 nm and the specific surface area of 29.2 m^2/g. For comparison, the pigments

prepared in a conventional batch reactor are significantly larger (mean particle size of 34 nm with a standard deviation of 20 nm) and the specific surface area (12.6 m^2/g) is only 40% of the above mentioned value.

The pigments Yellow 12, Red 2, Red 146, and Yellow 191 based on an azo-group were synthesized and compared with pigments that were produced conventionally [68]. The first three pigments are used as offset printing colors, whereas Yellow 191 was investigated as a dye for poly(vinylchlorid) (PVC) foils. The printing colors synthesized using the MicroJet Reactor have a superior color strength, brightness, and transparency compared to conventionally synthesized pigments. In the case of Yellow 12, it has been reported [68] that the mean particle size is about 30% reduced compared to the commercially available pigment.

The pigment Yellow 191 was used as a dye to color a PVC foil. Using the pigment prepared with a MicroJet Reactor, the color strength was increased and the color distribution was more homogenous compared to the use of the conventionally produced pigment Yellow 191.

In an exploratory series of experiments, $CaCO_3$ powder was synthesized [69] using $Ca(NO_2)_3$ ($c = 40$ mmol) and K_2CO_3 ($c = 40$ mmol) aqueous reaction solutions. The concentration of the surfactant TWEEN 80 was varied from 0, to 0.05, to 0.5%. The SEM picture (Fig. 7) shows the powder produced without surfactant. The particle size was about 12 μm. The particle size decreases ($d = 6$ μm and 3 μm) with increasing concentration ($c = 0.05\%$ and 0.5%) of the surfactant. In the case of the high surfactant concentration, particles with an additional type of shape were formed. In all cases, the particles look homogeneous in size. Significantly smaller or significantly larger particles were not observed.

8. CPC REACTOR

Laboratory scale CPC microreactors are assembled of microstructured plates mounted together by metallic bonding [70]. Each plate has numerous parallel channels in the submillimeter range with a special function such as dividing reactants into substreams and guiding them into the reaction channels, collecting the substreams with the product and leading it out of the reactor, as well as guiding the heat exchanger liquid to maintain isothermal reaction conditions. In the reaction zone, the flow is laminar, and the residence time is in the range of some seconds at a flow rate between 20 and 80 ml/min.

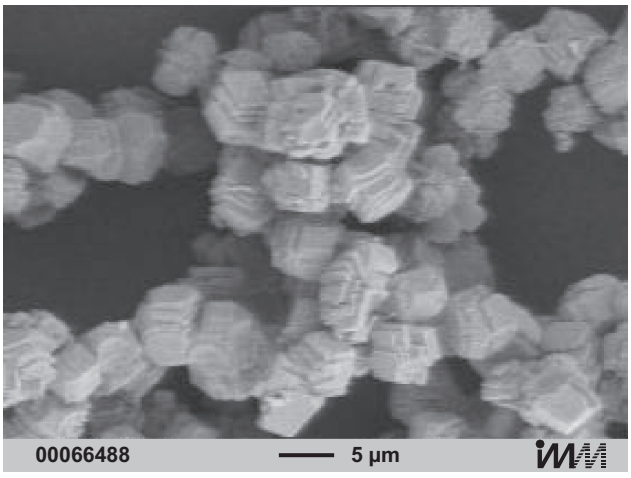


Figure 7. SEM picture of calcium carbonate produced with the MicroJet Reactor.

The pilot plant microreactor is based on a numbering-up concept. The reaction plate inside the microreactor was multiplied several times, and three of the reactors were integrated in one housing. Additionally, the dimensions of the reaction channels were slightly increased without changing the laminar flow regime or the mixing process. Also, in the pilot plant isothermal reaction conditions were maintained. With this pilot plant, a throughput of 500 ml/min reaction solution could be achieved.

Two different types of azo pigments were synthesized in a lab scale and pilot plant microreactor [71]. For pigment No. 2 the mean particle size (90 nm), is reduced drastically compared to the batch process (mean particle size: 600 nm), and a monomodal distribution (span of the particle size distribution = 1.5) is obtained instead of a bimodal distribution (span = 2) in the batch process. Beside these powder properties, the color properties could be dramatically improved as well. The color strength is increased by a factor of 1.49 using a microreactor compared to batch type synthesis. Additionally, the brightness and the transparency are enhanced by 6 steps on the glossier scale.

The estimated productivity of the synthesis of pigment No. 2 in the lab scale reactor is 1 t/a. To enhance this productivity to 10 t/a, a pilot plant, based on the numbering-up concept, was realized. This pilot plant runs several days successively with only cleaning steps of a few seconds every two or three hours having to be performed. The characterization of the pigments produced with the pilot plant shows the reproducibility of the powder characteristics found in the lab scale device.

9. SUMMARY

The production of powders using reactors equipped with microcomponents, named microreactors, is a rather new approach to obtain particles with defined morphology, small size, and, most important, narrow particle size distribution. Conventional stirred tank reactors suffer from mainly two disadvantages. First, the mixing of the reactants is ineffective and inhomogeneous, which results in a broad particle size distribution. Secondly, the scale-up procedure is rather risky because the powder quality is often different in a small-scale laboratory batch reactor compared to a large-scale production plant.

In contrast to the conventional reactors, all microreactors described here allow an intensive and, above all, homogenous mixing of the reactants. Furthermore, a continuous operation is achieved with a continuous flow of the reactant solution as well as a defined residence time of the precipitated powder can be performed. For production purposes, these operation conditions allow the numbering-up of the microreactor. This procedure avoids the risky scale-up step with all the drawbacks mentioned above.

Several examples of powders produced in different types of microreactors have been described. Ceramic powders such as barium titanate and calcium carbonate as well as

organic pigments are synthesized just to name a few. All together, these powders have a superior quality compared to batch-type synthesis. In most cases, small particles with a narrow particle size distribution as well as a defined morphology can be obtained. Additionally, organic pigments produced with a microreactor have superior color properties (such as brightness) compared to pigments produced in a conventional production process.

Despite the short time these microreactors have been used for powder synthesis, they are not only of scientific interest but also of industrial relevance. Pilot plants of different microreactors are currently being built and investigated. The first industrial production of powders with this new technology is expected to start in the near future [72]. These pilot plants will be able to produce several tons of powder per year in a reproducible high quality.

GLOSSARY

Impinging-jet micromixer Special type of micromixer where mixing is performed in a chamber outside the mixer ("chamber-free mixing").

Micromixer A special of microreactor for mixing purposes. In general, micromixers provide fast and homogeneous mixing.

Micropowder The mean particle size is in the range of 1 to 1000 μm. The powders dsecribed in this chapter have a size below 20 μm.

Microreactor Microreactor is used as a term for a chemical reactor of undefined size that contains components with microstructured channels or tubes in the microscale. In general, the width of such channels is in the micrometer range, mostly between 50 to 500 μm.

Nanopowder The mean particle size is in the range of 1 to 1000 nm.

Numbering-up For increasing the productivity of a production plant, numbering-up is a method of scaling by replication instead of the conventional scaling by dimensional changes, which is termed scale-up.

Pigment Solid particle of a dye.

Scanning electron microscopy (SEM) Technique for making enlarged pictures of objects with an electron beam. The resolution is enlarged compared to conventional light microscopy. Smaller objects can be seen.

Separation-layer micromixer Special type of micromixer were the two reaction solutions are separated by pure solvent ("delay-type mixing").

ACKNOWLEDGMENT

The European Community under the "Competitive and Sustainable Growth" program, contract No. G5RD-CT-1999-00123, is gratefully acknowledged for financial support.

REFERENCES

1. D. Ganguli and M. Chatterjee, "Ceramic Powder Preparation: A Handbook," p. 12, Kluwer Academic Publishers, Boston, 1997.
2. J. Falbe and M. Regitz, Eds., "Römpp Chemie Lexikon," 9th edition, p. 3680. Georg Thieme Publisher, Stuttgart, New York, 1992.
3. M. A. Hines and P. J. Guyot-Sionnest, *Phys. Chem.* 100, 468 (1996).
4. H. Scherzberg, K. Kahle, and K. Käseberg, *Chem. Eng. Technol.* 21, 412 (1998).
5. C. D. Rielly and A. J. Marquis, *Chem. Eng. Sci.* 56, 2475 (2001).
6. P. Bowen, N. Jongen, M. Donnet, J. Lemaître, H. Hofmann, R. Schenk, C. Hofmann, M. Aoun-Habbache, S. Guillemet-Fritsch, J. Sarrias, A. Rousset, M. Viviani, M. T. Buscaglia, V. Buscaglia, P. Nanni, A. Testino, and J. R. Herguijela, "Proceedings of the 4th World Congress on Particle Technology," Sydney, Australia, 2002.
7. T. Zech, D. Hönicke, M. Fichtner, and K. Schubert, "Proceedings of the 4th International Conference on Microreaction Technology," p. 390, Atlanta, GA, USA, 2000.
8. V. Hessel and H. Löwe, *Chem.-Ing.-Tech.* 74, 17 (2002).
9. T. Herweck, S. Hardt, V. Hessel, H. Löwe, C. Hofmann, F. Weise, T. Dietrich, and A. Feiertag, "Proceedings of 5th International Conference on Microreaction Technology," p. 215. Strasbourg, France, 2001.
10. C. Sangregorio, E. E. Carpenter, and C. J. O'Connor, *Proceedings of the Materials Research Society Symposium*, Vol. 577, 435 (1999).
11. G. X. Cheng, F. Shen, L. F. Yang, L. R. Ma, Y. Tang, K. De Yao, and P. C. Sun, *Mater. Chem. Phys.* 56, 97 (1998).
12. F. Gauffre and D. Roux, *Langmuir* 15, 3738 (1999).
13. K. Landfester, *Adv. Mater.* 13, 765 (2001).
14. B. Xia, I. W. Lenggero, and K. Okuyama, *Chem. Mater.* 14, 2623 (2002).
15. V. Hessel, H. Löwe, and T. Stange, *Lab on a Chip* 2, 14 (2002).
16. J. Brandner, M. Fichtner, K. Schubert, M. Liauw, and G. Emig, "Proceedings of the 5th International Conference on Microreaction Technology," p. 164. Strasbourg, France, 2001.
17. O. Wörz, K. P. Jäckel, T. Richter, and A. Wolf, "Proceedings of the 2nd International Conference on Microreaction Technology," New Orleans, LA, USA, 1998.
18. K. Schubert, *Chem.-Ing.-Tech.* 72, 24 (1998).
19. H. S. Bergh, J. R. Engstrom, S. Guan, A. Hagemeyer, C. Lugmair, H. Turner, L. van Erden, and W. H. Weinberg, "Proceedings of the 5th International Conference on Microreaction Technology," p. 191. Strasbourg, France, 2001.
20. A. Rouge, B. Spoetzl, K. Gebauer, R. Schenk, and A. Renken, *Chem. Eng. Sci.* 56, 1419 (2001).
21. A. Müller, K. Drese, H. Gnaser, M. Hampe, V. Hessel, H. Löwe, S. Schmitt, and R. Zapf, *Catal. Today* 81, 377 (2003).
22. H. Lu, M. A. Schmidt, and K. F. Jensen, "Proceedings of the 5th International Conference on Microreaction Technology," p. 175. Strasbourg, France, 2001.
23. H. Lu, M. A. Schmidt, and K. F. Jensen, *Lab on a Chip* 1, 22 (2001).
24. H. Löwe, M. Küpper, and A. Ziogas, German Patent Application, DE 19841302 (1998).
25. A. Ziogas, H. Löwe, M. Küpper, and W. Ehrfeld, in "Microreaction Technology: Industrial Processes" (W. Ehrfeld, Ed.), p. 136. Springer, Berlin, 1998.
26. M. Küpper, V. Hessel, and H. Löwe, *Electrochim. Acta* 48, 2889 (2003).
27. G. Veser, G. Friedrich, M. Freygang, and R. Zengerle, in "Microreaction Technology: Industrial Processes" (W. Ehrfeld, Ed.), p. 674. Springer, Berlin, 1998.
28. V. Hessel and H. Löwe, *Chem.-Ing.-Techn.* 74, 185 (2002).
29. K. Jähnisch, M. Baerns, V. Hessel, W. Ehrfeld, V. Haverkamp, H. Löwe, C. Wille, and A. Guber, *J. Fluorine Chem.* 105, 117 (2000).
30. Y.-S. Her, E. Matijevic, and W. R. Wilcox, *Powder Technol.* 61, 172 (1990).
31. Y.-S. Her, S.-H. Lee, and E. Matijevic, *J. Mater. Res.* 11, 156 (1996).
32. K. Y. Park, M. Ullmann, Y. J. Suh, and S. K. Friedlander, *Journal of Nanoparticle Research* 3, 309 (2001).
33. F. Kirkbir and H. Komiyama, *Can. J. Chem. Eng.* 65, 759 (1987).
34. J. B. Edel, R. Fortt, J. C. deMello, and A. J. de Mello, *Chem. Commun.* 1136 (2002).

35. F. G. Bessoth, A. J. deMello, and A. Manz, *Anal. Commun.* 36, 213 (1999).
36. J. Lemaître, N. Jongen, R. Vacassy, and P. Bowen, Patent, WO 98/02237 (1998); N. Jongen, J. Lemaître, P. Bowen, and H. Hofmann, "Proceedings of the 5th World Congress of Chemical Engineering," San Diego, CA, USA, Vol. V, 31 (1996).
37. P. Penicot, H. Muhr, E. Plasari, and J. Villermaux, *Chem. Eng. Technol.* 21, 507 (1998).
38. R. David and B. Marcant, *AIChE J.* 40, 424 (1994).
39. B. Werner, M. Donnet, V. Hessel, C. Hofmann, N. Jongen, H. Löwe, R. Schenk, and A. Ziogas, "Proceedings of the 6th International Conference on Microreaction Technology," p. 168, New Orleans, CA, USA, 2002.
40. R. Schenk, V. Hessel, C. Hofmann, F. Schönfeld, and B. Werner, German Patent Application, DE 101 48 615.4 (2001).
41. R. Schenk, V. Hessel, B. Werner, F. Schönfeld, Ch. Hofmann, M. Donnet, and N. Jongen, *Chemical Engineering Transactions* 1, 909 (2002).
42. J. Villermaux, L. Falk, M.-C. Fournier, and C. Detrez, *AIChE Symp. Ser.* 286, 6 (1991).
43. P.-H. Jézéquel, *Chem. Eng. Sci.* 56, 2399 (2001).
44. N. Jongen, M. Donnet, P. Bowen, J. Lemaître, H. Hofmann, R. Schenk, C. Hofmann, M. Aoun-Habbache, S. Guillemet-Fritsch, J. Sarrias, A. Rousset, M. Viviani, M. T. Buscaglia, V. Buscaglia, P. Nanni, A. Testino, and J. R. Herguijela, *Chemical Engineering Transactions* 1, 807 (2002).
45. B. Marcant, Thesis, Institut National Polytechnique de Lovaine, France, 1992.
46. M. Donnet, N. Jongen, J. Lemaître, P. Bowen, and H. Hofmann, "Proceedings of the 14th Symposium on Industrial Crystallization," CD, 82, Institution of Chemical Engineers, Rugby, Cambridge, UK, 1999.
47. M. Donnet, N. Jongen, J. Lemaître, and P. Bowen, *J. Mater. Sci. Lett.* 19, 749 (2000).
48. L. Brecevic, D. Kralj, and J. Garside, *J. Cryst. Growth* 97, 460 (1989).
49. O. Söhnel and F. Grases, *Adv. Colloid Interface Sci.* 59, 1 (1995).
50. N. Jongen, J. Lemaître, P. Bowen, and H. Hofmann, *Chem. Mater.* 11, 712 (1999).
51. R. Vacassy, N. Jongen, J. Lemaître, P. Bowen, and H. Hofmann, "Proceedings of the 3rd International Particle Technology Forum," 375, Institution of Chemical Engineers, Rugby, Brighton, UK, July 6–9, 1998.
52. V. Hessel, W. Ehrfeld, H. Löwe, J. Schiewe, and M. Donnet, "Proceedings of 3rd International Conference on Micro Materials," p. 317. Berlin, 2000.
53. R. Vacassy, J. Lemaître, P. Bowen, H. Hofmann, and J. H. Gerlings, *AIChE J.* 46, 1241 (2000).
54. J. M. Marentette, J. Norwig, E. Stöckelmann, W. H. Meyer, and G. Wegner, *Adv. Mater.* 9, 647 (1997).
55. R. Schenk, M. Donnet, V. Hessel, C. Hofmann, N. Jongen, and H. Löwe, "Proceedings of the 5th International Conference on Microreaction Technology," p. 489. Strasbourg, France, 2001.
56. M. Donnet, P. Bowen, N. Jongen, J. Lemaître, H. Hofmann, A. Schreiner, A. G. Jones, R. Schenk, C. Hofmann, and S. De Carlo, *Chemical Engineering Transactions* 1, 1353 (2002).
57. A. Schreiner and A. G. Jones, *Chemical Engineering Transactions* 1, 1245 (2002).
58. M. Aoun-Habbache, S. Guillemet-Fritsch, J. Sarrias, A. Rousset, N. Jongen, M. Donnet, P. Bowen, and J. Lemaître, "Proceedings of the 10th International Ceramic Congress," Florence, Italy, 2002.
59. M. Aoun-Habbache, S. Guillemet-Fritsch, J. Sarrias, L. Godet-Morand, A. Rousset, P. Bowen, and N. Jongen, *Chemical Engineering Transactions* 1, 1281 (2002).
60. W. J. Sarjeant, J. Zirnheld, F. W. MacDougall, J. S. Bowers, N. Clark, I. W. Clelland, R. A. Price, M. Hudis, I. Kohlberg, G. McDuff, I. McNab, S. G. Parler, Jr., and J. Prymak, in "Handbook of Low and High Dielectric Constant Materials and Their Applications" (H. S. Nalwa, Ed.), Vol. 2, p. 423. Academic Press, San Diego, 1999.
61. B. Huybrechts, K. Ishizaki, and M. Takata, *J. Mater. Sci.* 30, 2463 (1995).
62. P. Bowen, M. Donnet, A. Testino, M. Viviani, M. T. Buscaglia, V. Buscaglia, and P. Nanni, *Key Eng. Mater.* 206–213, 21 (2002).
63. M. T. Buscaglia, V. Buscaglia, M. Viviani, A. Testino, P. Nanni, P. Bowen, M. Donnet, N. Jongen, R. Schenk, C. Hofmann, V. Hessel, and F. Schönfeld, "Proceedings of the 10th International Ceramic Congress," p. 535. Florence, Italy, 2002.
64. M. T. Buscaglia, V. Buscaglia, M. Viviani, A. Testino, P. Nanni, P. Bowen, M. Donnet, N. Jongen, R. Schenk, V. Hessel, and F. Schönfeld, *Chemical Engineering Transactions* 1, 1341 (2002).
65. B. Penth, *Process* 7, 62 (2000).
66. B. Penth, International Patent Application, WO 00/61275 (2000).
67. E. Dietz, J. Weber, D. Schnaitmann, Ch. Wille, L. Unverdorben, and K. Brychcy, European Patent Application, EP 1 195 413 A1 (2001).
68. E. Dietz, J. Weber, D. Schnaitmann, Ch. Wille, L. Unverdorben, K. Saltmacher, J. Jung, and K. Schiffer, European Patent Application, EP 1 195 411 A1 (2001).
69. R. Schenk, B. Werner, and B. Penth, unpublished results, 2002.
70. V. Autze, A. Kleemann, and S. Oberbeck, *Nachrichten aus der Chemie* 48, 683 (2000).
71. Ch. Wille, V. Autze, H. Kim, U. Nickel, S. Oberbeck, Th. Schwalbe, and L. Unverdorben, "Proceedings of the 6th International Conference on Microreaction Technology," New Orleans, CA USA, 2002.
72. Geschäftsbericht (annual report), Company Clariant, pp. 31–33 (2001).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanoprecipitates and Nanocavities in Functional Materials

J. Th. M. De Hosson

University of Groningen, Groningen, The Netherlands

A. van Veen

Delft University of Technology, Delft, The Netherlands

CONTENTS

1. INTRODUCTION

Nanoprecipitates and nanocavities can improve or degrade functional properties to a large extent. In particular nanocavities in low-k dielectrics affect the electronic properties of semiconductors that are employed in digital equipment [1]. In recent years a new era has emerged to control these defect clusters to provide novel functional properties. Examples are the metallic and semiconductor nanoclusters in optical materials, which enable one to tailor optical properties of the materials, with important applications, for example, in fiber optics. In silicon helium ion implantation induced nanocavities are used to getter metallic impurities in the design of devices, which does not allow one to use conventional gettering methods. This chapter aims at showing to what level this field of the design of nanocavities and nanoprecipitates has progressed. In all cases the precipitates are embedded in a matrix. This can be either a crystalline or an amorphous structure. The properties of the embedded precipitate or nanocluster depend for an important part on its surroundings, in particular on the interface between particle and the matrix. Furthermore, the shape and even the phase of the precipitate are determined by the surrounding medium. In the case of a crystalline matrix orientation relationships can exist, which cause crystalline alignment of the nanoclusters. This is in contrast to the other class of nanoparticles and assemblies of nanoparticles, which are grown in a liquid or gaseous medium and deposited as a layer on a substrate. In general, chemical growth modes give rather precise control over size. Implantation produced nanoparticles, which might be called a physical synthesis method, lack a narrow size distribution. However, certain tools are developed and will be developed in the future to narrow the size distribution or even create monodisperse particles. Also it has been suggested that a two-step process, where nanocavities are used as templates to nucleate and promote precipitate growth, offers new possibilities to create new types of precipitates. In principle, any atomic species can be implanted in any material. The starting point is displaced atoms, vacant lattice sites, and implanted atoms. Thermal annealing is used to remove the defects and promote clustering of the implanted atoms. Alloys and compound precipitates can be formed by multiple ion species implantation. A requirement is that all implanted species are poorly soluble in the matrix. Otherwise a homogeneous nucleation of the clusters will not occur.

Besides implantation methods, other methods exist, some of which have been known for more than 2000 years, to create metallic or other species in supersaturation in glassy materials. The methods consist of melting and quenching or reduction. Usually, these are bulk methods and graded layers cannot easily be produced. In this chapter we will restrict ourselves mainly to nanoprecipitates and nanocavities produced by ion implantation in crystalline media, notably single crystals of metal oxides and silicon. Production and optical properties of implanted metallic

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (297–349)

and semiconducting particles in amorphous SiO_2 have been reviewed by Polman [2], and by Schultz and Lynn [3].

Besides well known techniques suitable to observe precipitates and nanocavities, like transmission electron microscopy, including all the analytical tools added to the electron microscope in the last decade, the less known positron annihilation techniques have been introduced to monitor precipitates and cavities. A positron is a positively charged particle that is repelled by nuclei. As a result it has a high preference for open volume. Because of the preference of positrons for open volume, that is to say, much more sensitive than electrons in electron microscopy, positron annihilation spectroscopy (PAS) has been developed over the last few decades into a successful nondestructive method for probing low atomic density regions (e.g., vacancies, clusters of vacancies, and nanocavities) in materials over a wide range of depths, from the surface to depths of hundreds of nm [3]. PAS can thus provide information that is complementary to the results provided by many other techniques for the examination of the chemical and structural properties of materials. Actually, a wealth of techniques is available for these latter purposes. Among them are Rutherford backscattering, nuclear reaction analysis, and secondary ion mass spectrometry, all of which are capable of depth profiling of impurity atoms [4]. Nevertheless, these techniques are not sensitive to the presence of open-volume defects. An indirect method for probing open volume is small angle neutron or X-ray scattering. Open volume is best probed by using probe particles that are trapped in the cavities. Suitable particles are gas atoms such as helium and hydrogen, or positrons. Other techniques using a probe atom are Mössbauer and perturbed angular correlation, but these are better suited for probing very small cavities (1–5 vacancies). The use of positrons as a probe has the advantage that the material is not damaged by the irradiation with the probe particles. A review on defect profiling by positrons and other techniques has been given by Dupasquier and Ottaviani [5]. Furthermore, PAS has been extensively applied to the study of thin films, layers of embedded nanoclusters, and interfaces [6]. In this chapter we confine ourselves to positron annihilation spectroscopy as a complementary technique to highresolution transmission electron microscopy.

In recent years, the development of radioisotope based monoenergetic slow positron (e^+) beams has led to the study of interesting surface and near surface phenomena in thin films and ion implanted materials [7]. By varying the initial energy of the implanted positrons one can study defects at different depths. The importance of depth-profiling studies becomes clear if one considers layered systems obtained by various deposition or implantation techniques. In these systems defects may appear at the interface of the layers because of low mobility of the deposited atoms. In some cases the layer becomes amorphous and will show many defects. The relatively low intensity of the radioisotope based slow beams was offset by the use of efficient solid-state detectors. The combination of depth profiling with the high-resolution two-dimensional angular correlation of annihilation radiation (2D-ACAR) method for the study of surfaces and layered systems, however, had to wait for the development of high-intensity slow positron beams. Nowadays, such beams are operational or under construction in several laboratories [8]. Depth profiling positron 2D-ACAR thin film studies have been performed for the first time at Brookhaven National Laboratory on the Si/SiO_2 system employing a strong ^{64}Cu source produced in a high-flux beam reactor [9]. The recent construction of a stable high-intensity reactor-based slow positron beam at the Delft 2 MW research reactor has proven to be successful [10, 11]. The beam called POSH ("POSitrons at the Hoger onderwijs reactor") has an intensity of $\sim 4 \times 10^8$ e^+/s. Numerous depth-profiling 2D-ACAR studies of various materials of great interest in materials science have been performed, (e.g., nanocavities in Si and MgO produced by ion implantation, metallic nanoprecipitates in MgO, SIMOX, low-k dielectric and silicon carbide thin films) [12–18].

The structure of this chapter is as follows. Section 2 presents a general introduction into positron physics, positron annihilation techniques, and intense positron sources. Further, some essential features of high-resolution electron microscopy are also presented. The new possibilities have been successfully applied to various types of materials, the results of which are collected in the subsequent Sections 3–7.

2. THEORETICAL BACKGROUND OF THE EXPERIMENTAL TECHNIQUES

2.1. Positron

The positron is the antiparticle of the electron. Hence, it has the same mass and the same spin (1/2), but opposite charge and magnetic moment. Dirac postulated the positron in 1930 as the "negative" energy extension of his theory of electron energy levels [19, 20]. Two years later Anderson [21] discovered the positron in his cloud chamber experiments. Positrons are unstable in matter where they annihilate with electrons predominantly via 2γ decay, both photons carrying energy of 511 keV. In rare circumstances (e.g., when a positron is in the vicinity of a heavy nucleus) the positron and the electron can undergo a 1γ decay process. The emission of more photons is also possible, but the probability of such an event is small (the ratio between the probabilities of 3γ and 2γ decay in metals is 1/371 [22]).

When a positron is injected into a material it interacts with its surroundings and rapidly loses its energy by ionizing collisions with atoms, collisions with conducting electrons, and phonon scattering, until it reaches thermal equilibrium. This process occurs within 10^{-11} s, which is short compared to the mean lifetime of the positron of the order of 10^{-9}–10^{-10} s [23]. This lifetime is so short that for realistic source strengths no more than one positron is present in the sample at any time. At room temperature the wavelength associated with a thermalized positron is usually one order of magnitude larger than the typical lattice parameter. Hence, in a solid the positron behaves like a delocalized Bloch wave, which in well-annealed crystals of metals and semiconductors exhibits maxima in the interstitial region due to core repulsion. In defective materials, however, the positron can be localized (trapped) in, or around, low-density regions such as vacancies, dislocations, and microvoids. As a consequence, the positron lifetime becomes longer and the

positron density is strongly enhanced at the position of a defect [24].

Positrons can be produced by the nuclear decay of isotopes or by pair production by high-energy gamma rays ($E_\gamma > 1.022$ MeV). Positrons emitted by a radioactive isotope (e.g., ^{22}Na, ^{58}Co, ^{64}Cu) have continuous energy distributions ranging from 0 to a few hundred keV. As a consequence, when positrons are implanted in a material, the implantation profile is broad, of the order of a few hundred micrometers. Because positrons are positively charged particles, they are repelled by the atomic cores, thus leading to the possibility of their ejection from a surface due to their negative work function ($\phi_+ < 0$). The positron work function for a solid is defined as the minimum energy required to remove a positron from a point inside the solid to one in the vacuum [3]. Owing to its interaction with the surrounding electrons, a positron can adopt an energy level that can be higher or lower than the energy level of the vacuum. As a consequence, the positron work function can be negative or positive in comparison to the electron work function, of which the latter is nearly always negative. The effect of positron ejection from a surface is used in the production of slow positron beams. Examples of surfaces that emit positrons are clean metal surfaces, for example, W(111), Ni(100), Cu(111), diamond, Si, SiC, GaN, or the surfaces of rare-gas solids.

2.1.1. Positronium

In some liquids and insulating solids, the positron can capture an electron prior to annihilation, forming positronium (Ps), a hydrogenlike positron–electron bound state with a binding energy in vacuum of ~6.8 eV. Predicted classically by Mohorovicic [25] in 1934 and quantum mechanically by Ruark [26] in 1945, the Ps atom was experimentally discovered in 1951 by Deutsch who found a long-lived component in a lifetime spectrum that was nearly independent of gas pressure [27].

The spectroscopic differences between Ps and hydrogen originate in the particle–antiparticle nature of Ps (equality of the positron and electron masses and of the magnitudes of their magnetic moments), which lead to the possibility of self-annihilation. The ground states of Ps consist of a singlet $^1S_0$ state ($S = 0; m_s = 0$) with the spins oriented antiparallel, called para-positronium (p-Ps), and a triplet $^3S_1$ state ($S = 1; m_s = 0, \pm 1$) with the spins oriented parallel, called ortho-positronium (o-Ps). Para-Ps has a lifetime in vacuum of ~125 ps and decays via two γ rays, while o-Ps has a lifetime in vacuum of ~142 ns decaying into three or more γ rays. In the absence of external electric and magnetic fields the relative amount of p-Ps:o-Ps formed is 1:3. In a low density gas the long-lifetime component approaches that characteristic of o-Ps, while at higher densities or in condensed matter the interaction of Ps with atomic electrons leads to a decrease in the 3γ decay mode associated with a shortening of the long lifetime. This decrease is due to the so-called "pick-off" process in which the positron of o-Ps annihilates during a collision with an atomic electron that has an antiparallel spin and thus annihilates via a 2γ process.

2.1.2. Magnetic Field Induced 2γ Annihilation

In the presence of an external magnetic field of flux density B the p-Ps state ($^1\Phi_0$) and the $m_s = 0$ substate of o-Ps ($^3\Phi_0$) are mixed in so-called ortho-like-Ps and para-like-Ps states [28–30],

$$\Phi_+ = \frac{1}{\sqrt{1+y^2}}{}^3\Phi_0 + \frac{y}{\sqrt{1+y^2}}{}^1\Phi_0 \tag{1}$$

$$\Phi_- = \frac{-y}{\sqrt{1+y^2}}{}^3\Phi_0 + \frac{1}{\sqrt{1+y^2}}{}^1\Phi_0 \tag{2}$$

with

$$y = \frac{x}{\sqrt{1+x^2}+1} \qquad \text{and} \qquad x = \frac{4\mu_B B}{\Delta E} \tag{3}$$

Here, μ_B is the Bohr magnetron and $\Delta E = 8.4 \times 10^{-4}$ eV is the hyperfine splitting between the p-Ps state and the o-Ps states. The annihilating rates of the new states are

$$\Gamma_+ = \frac{1}{1+y^2}\Gamma_3 + \frac{y^2}{1+y^2}\Gamma_1 \tag{4}$$

for ortho-like-Ps and

$$\Gamma_- = \frac{y^2}{1+y^2}\Gamma_3 + \frac{1}{1+y^2}\Gamma_1 \tag{5}$$

for para-like-Ps, where $\Gamma_3 = (142\ \text{ns})^{-1}$ and $\Gamma_1 = (125\ \text{ps})^{-1}$ are the annihilation rates of o-Ps and p-Ps, respectively. Both these new states can undergo 2γ and 3γ annihilations. Since Γ_1 is three orders of magnitude larger than Γ_3, even a weak magnetic field for which $y \ll 1$ causes a large fraction of ortho-like-Ps atoms to self-annihilate into 2γ. The annihilation characteristic of para-like-Ps for small y, however, is not so much affected by the magnetic field. This effect is called magnetic field induced o-Ps $\rightarrow$ p-Ps conversion or magnetic quenching of o-Ps.

Positronium annihilates subsequently with a characteristic signature and can be detected by studying the annihilation γ's in different manners: (i) by measuring the 3γ to 2γ yield ratio with a multiple-coincidence γ detector system, (ii) by measuring the positron lifetime, (iii) by measuring the energy spectrum of one of the annihilation photons, (iv) by measuring the momentum distribution of the electron–positron pairs, and (v) by studying the decrease in the 3γ yield due to the Zeeman mixing of Ps states when a magnetic field is applied [31–33]. In this chapter the attention will be focused on (iv).

Positronium is usually formed in regions with a low electron density such as gases or liquids. In gases Ps formation takes place by radiationless electron capture as long as the kinetic energy (T) of the positron is in the energy region $E_i - 6.8 \le T \le E_i$ [eV], called the Ore gap, where E_i is the ionization energy of the gas atom or molecule [22]. Ps can also be formed in some insulating materials and organic materials with large open volumes where an electron is captured in the ionization trail ("spur") created by the slowing down of the positron [34]. In these materials p-Ps still has a lifetime of 125 ps, but the lifetime of o-Ps is reduced to

a few nanoseconds owing to pick-off processes. Since the Pauli exclusion principle is not operative for a Ps atom (Ps is a boson), the Ps atom becomes quickly thermalized and thus resides near the bottom of its own band as a virtually momentum-free system. As a consequence, self-annihilation of Ps contributes a sharp peak in a 2D-ACAR spectrum. In metals and semiconductors, thermalized positrons annihilate from delocalized Bloch states without forming Ps.

With respect to the spatial extension of the Ps wave function (Ψ_{Ps}), two different types of Ps can be distinguished:

(i) *Localized Ps*, when $\Psi_{Ps}(\mathbf{r})$ extends over a region containing no, or only a few, atoms of the medium. Figure 1a shows the experimental 2D-ACAR momentum distribution of localized *p*-Ps atoms formed in radiation produced microvoids in vanadium.

(ii) *Delocalized Ps*, when $\Psi_{Ps}(\mathbf{r})$ extends over a region containing a large number of elementary cells of a crystalline solid. Figure 1b shows the experimental 2D-ACAR momentum distribution of delocalized *p*-Ps formed in an oriented crystal of SiO_2.

In some circumstances (e.g., when Ps is trapped in a large void) the positronium atom formed cannot lose its energy by collisions with atoms and molecules, thus remaining nonthermalized or "hot" Ps [12]. Also, when positrons are implanted at shallow depths in metals or semiconductors, emission of energetic Ps atom from the surface may occur [35, 36].

2.2. Positron Annihilation Spectroscopies

The positron annihilation process possesses a number of signatures that contain information concerning the system under investigation. Various experimental methods have been developed during the past decades to measure these signatures. In this section we shall discuss how and which information can be extracted from the experimental data.

There are three major positron techniques that use the annihilation photons in order to gain information about the electrons: (a) two-dimensional angular correlation of annihilation radiation, (b) Doppler broadening of annihilation radiation experiments, and (c) positron lifetime experiments. The first two techniques give information on the momentum distribution of electrons while the third one gives information on the electron density. Other techniques use different properties of the positron to probe material properties, (e.g., low-energy positron diffraction, positron channeling, positron annihilation induced Auger electron spectroscopy, but they are not the subject of this chapter [3, 37].

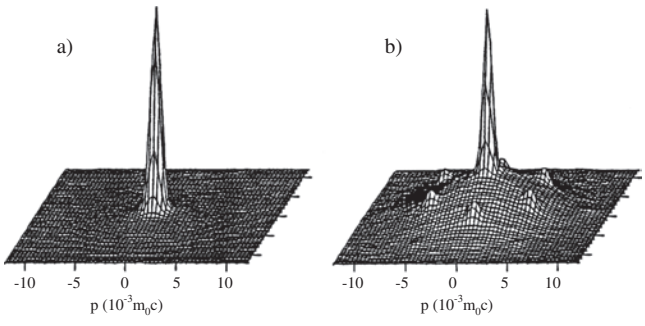


Figure 1. (a) Positronium localized in radiation produced microvoids in vanadium. (b) Delocalized positronium in an oriented crystal of SiO_2. Reprinted with permission from [5], A. Dupasquier and G. Uttaviani, in "Position Spectroscopy of Solids" (A. Dupasquier and A. P. Mills, Jr., Eds.), pp. 581–657. IOS, Amsterdam, 1995. © 1995, IOS.

A thermalized positron and an electron annihilate under emission of mostly two γ quanta. In the center-of-mass system of the annihilating particles the γ quanta are emitted in exactly opposite directions and the available energy of $2m_0c^2 - E_B$ (where $E_B = E_B^- + E_B^+$ denotes the binding energy of the two particles in the solid) is divided equally over the two quanta. In the laboratory coordinate system the momentum $\boldsymbol{p}$ of the two particles (largely determined by the electron, the positron being thermal) manifests itself in two different ways. First, the momentum component $p_{||}$ along the direction of γ-ray emission results in an energy shift $\Delta E = \pm 1/2\ p_{||}c$ of the energy of the γ quantum from its average value $m_0c^2 - 1/2E_B$. Since one half of the electrons moves toward the detector while the other half moves away from it, this shift results in a Doppler broadening by several keVs of the 511 keV line which can be measured with a solid-state detector [38]. Since the distance between the detector and the sample is usually very small the Doppler broadening experiment integrates over virtually all values of the transverse component of the momentum.

This transverse component can be measured (in one or two dimensions) in 1D- or 2D-ACAR experiment [39, 40]. In the laboratory system of reference, the directions of γ-ray emission differ from 180° by a small angle $\theta \sim p_\perp/m_0c$, typically of the order of 10 mrad and proportional to the component $p_\perp$ of the momentum perpendicular to the direction of γ-ray emission. In view of the smallness of θ the latter type of experiment requires large sample–detector separations in order to realize a good momentum resolution, and thus a strong radioisotope source or an intense positron beam is needed.

A third signature of the annihilation process is the positron lifetime, which is the inverse of the rate at which positrons annihilate. The lifetime depends on the density of electrons at the position of the positron. It can be measured in a delayed coincidence experiment in which the time interval is measured between the emission of a positron by the positron source and the entrance of a positron into the sample, and its subsequent annihilation. Finally, positrons diffusing back to the surface or trapped at the surface of the sample may be reemitted after having formed positronium, the hydrogenlike bound state of a positron and an electron. This so-called Ps fraction is usually not measured directly but derived in an indirect way from other measurements. In the following we shall examine the various signatures of the annihilation process more closely.

2.2.1. Two-Dimensional ACAR

The measurement of the angular correlation of the annihilation radiation is one of the means for studying positron annihilation in solids. This technique measures the angular deviation from collinearity of the two coincident annihilation photons. In the center-of-mass frame the total momentum

of the e^+–e^- pair is zero and since the energy and momentum before and after annihilation are conserved the two photons are emitted in opposite directions, each one having an energy equal to m_0c^2 (see Fig. 2a). In the laboratory frame, however, the e^+–e^- pair carries a total momentum **p**. Since the positron is thermal and annihilates mainly with weakly bound electrons, the momentum **p** is relatively small ($\sim 10^{-2}\ m_0c$) and, as a consequence, it will slightly change the picture found in the center-of-mass frame. Hence, the components of the momentum perpendicular to the photon emission axis, p_x and p_y, will generate small angular deviations from collinearity (see Fig. 2b) that can be written as follows

$$\theta \sim \sin\theta = \frac{p_x}{m_0c} \tag{6}$$

$$\phi \sim \sin\phi = \frac{p_y}{m_0c} \tag{7}$$

Figure 3 shows a schematic diagram of a typical 2D-ACAR setup. It consists of two large-area position sensitive gamma detectors, a central source–sample vacuum chamber, and data acquisition electronics. In the central chamber a positron beam is guided by an axial magnetic field to the sample of interest. The annihilation quanta are detected in the two detectors and the information concerning their individual points of arrival is converted by the acquisition electronics into (θ, ϕ) addresses, which are stored in an incrementally updated 2D discrete array. In this way one obtains the two-dimensional projection of the three-dimensional momentum density $\rho^{2\gamma}(\mathbf{p})$ [41],

$$N(\theta,\phi) = N(p_x, p_y) = \left(\int \rho^{2\gamma}(\mathbf{p})\, dp_z\right) \otimes R(p_x, p_y) \tag{8}$$

where p_x, p_y, p_z are the components of the total momentum **p** (p_z is not measured) and $\otimes R$ denotes the convolution with the 2D angular resolution $R(p_x, p_y)$ of the setup.

Typical resolutions obtainable are of the order of 1 mrad. The samples are usually oriented with one of their main crystallographic directions parallel to the integration axis of the setup. Figure 4 shows an example of a 2D-ACAR spectrum

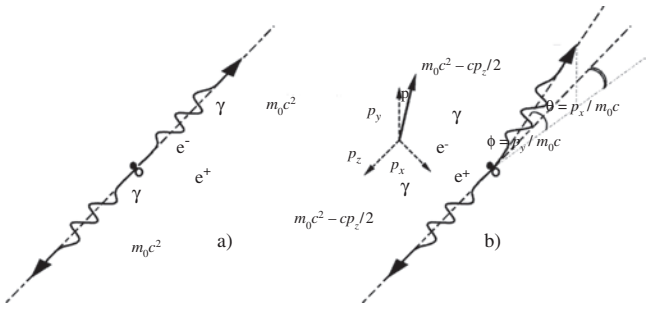


Figure 2. Two-photon annihilation process in: (a) center-of-mass frame; (b) laboratory frame.

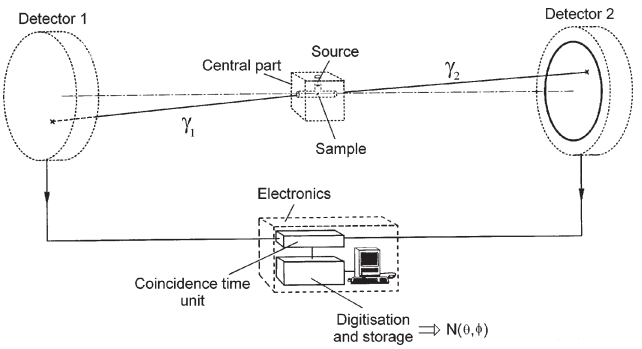


Figure 3. Schematic diagram of a typical 2D-ACAR setup showing the principal components.

for an oriented crystal of Si(100). The structure in the spectrum reflects the crystalline features of the momentum density of the electrons in Si. In the older 1D-ACAR method only one component of the total momentum was measured,

$$N(p_z) \sim \left(\iint \rho^{2\gamma}(\mathbf{p})\, dp_x\, dp_y\right) \otimes R(p_z) \tag{9}$$

where $\otimes R$ corresponds to the convolution with the one-dimensional angular resolution $R(p_z)$ of the setup. Since the 2D-ACAR method maps the momentum distribution of the electrons with a high-resolution, the method has important applications (e.g., Fermi surface topology).

The pioneering 2D-ACAR work with beams was performed in the 1980s in the laboratories of Lawrence Livermore (LLNL) and Brookhaven (BNL). The Livermore beam was made by moderation of high-energy positrons produced by pair production in a W target at the electron linac. The positrons were then moderated in a system of W vanes. The Brookhaven group used a slow positron source made by evaporating ~20 Ci of ^{64}Cu obtained by irradiation in the High-Flux Beam Reactor onto a single-crystal W(110) substrate, where it was annealed to form an epitaxial Cu(111) layer. This thin Cu layer emitted and simultaneously moderated the positrons into a monoenergetic beam of slow positrons. Both beams could be accelerated to obtain the desired incident positron energy.

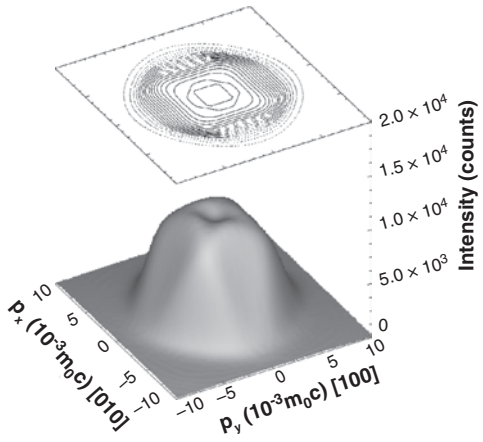


Figure 4. 2D-ACAR spectrum for Si (100).

These first 2D-ACAR experiments concentrated on the differences between the bulk and surface momentum distributions of single-crystal Al (BNL) [42–44] and Cu (LLNL) [45] samples and observed the momentum distribution of Ps emitted by the surface.

2.2.2. Doppler Broadening of Annihilation Radiation

In addition to the angular deviation of the annihilation photons, discussed in the previous section, there is a shift of the energy of the photons from its average value of 511 keV, caused by the component of the momentum parallel to the direction of γ-ray emission. The energy shift, δE, from 511 keV can be expressed by

$$\delta E = cp_z/2 \tag{10}$$

where p_z is the longitudinal component of the momentum of the e^+–e^- pair. Since one half of the electrons moves toward the detector while the other half moves away from it, the energy shift results in a Doppler broadening of the 511 keV line by several keVs, which can be measured by Ge detectors. The best detectors have a resolution of 1.1 keV at 511 keV which corresponds to a momentum of $4.3 \times 10^{-3}\ m_0c$. Owing to the small distance between the detector and the sample, the Doppler broadening experiment integrates over the transverse component of the momentum and the measured spectrum is a one-dimensional projection of the three-dimensional momentum density,

$$N(E) = N(p_z) \sim \left(\iint \rho^{2\gamma}(\mathbf{p})\, dp_x\ dp_y \right) \otimes R \tag{11}$$

where $\otimes R$ denotes convolution with the energy resolution of the Ge detector.

Figure 5 shows a typical Doppler broadening spectrum consisting of a broadened 511 keV peak. Due to the double integration and convolution with the resolution of the detectors the spectrum is smooth and featureless. However, one can derive valuable information from this spectrum by defining two regions of interest: the low-momentum central part corresponding to annihilations of positrons with valence electrons, and the high-momentum tails corresponding to annihilations with more tightly bound electrons (e.g., core electrons). Thus, one can characterize the shape of the 511 keV peak with the aid of two parameters, S and W, defined as [46, 47]

$$S = \frac{B}{A + B + C} \tag{12}$$

$$W = D + E + F + (F/D) \tag{13}$$

where A, B, C, D, E, and F are the areas of the energy windows defined in Figure 5. Therefore, the S parameter is more useful for monitoring the presence of defects in a material, while the W parameter gives the signature of that material. The parameter S gives the area under the central part of the profile, divided by the total area under the profile; thus a high value of S signals the presence of open-volume defects, whereas a "defect-free" sample will show a low value of the S parameter. The energy windows are set symmetrically around the center of the peak in such a way as to maximize the sensitivity to vacancies. Many workers use an S value of ~0.50. In order to make S parameter measurements in different laboratories easier to compare, the S values are usually quoted with respect to some standard material, often a highly perfect Si crystal. The orientation of this reference crystal and that of the crystal under investigation, if the latter also is a single crystal, should be specified, as the S parameter is an anisotropic quantity with the point symmetry of the crystal. The second parameter, W, gives the ratio between the area under the wings and the total area. The settings of the energy windows, which define the W parameter, differ more among different laboratories. Some workers choose their settings such that the value of W is approximately 0.25 of that of S, thereby obtaining a relatively high statistical precision, while others choose a lower fraction ($\leq$0.1) to increase the sensitivity to defects. As *areas under parts of the Doppler broadening profile give S and W parameters* they can be measured within a short time with a good statistical precision without putting high demands on the positron beam intensity. This makes this type of measurement highly suitable for defect investigations in which some parameter such as the positron energy, sample temperature, or previous heat treatment is varied.

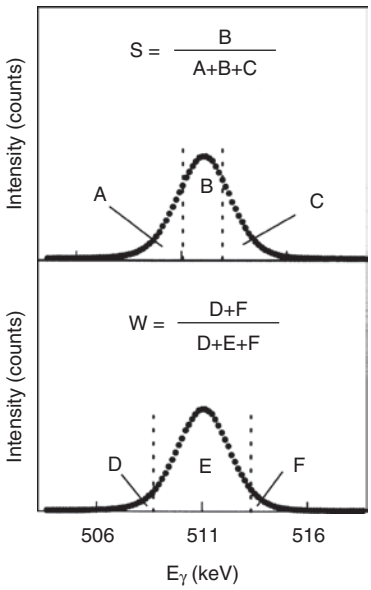


Figure 5. Definition of the S and W parameters in Doppler broadening spectrum.

In view of the positron affinity for open volumes, the Doppler broadening of the annihilation radiation technique in combination with slow positron beams has been extensively applied to the depth profiling of defects in materials [3, 6]. In these experiments the S and W parameters are measured as a function of the incident positron energy, E. The implantation of a slow positron beam can be approximated by a Makhovian profile [48],

$$P(z, E) = \frac{mz^{m-1}}{z_0^m} \exp\left[-\left(\frac{z}{z_0} \right)^m \right] \tag{14}$$

where m is a shape parameter usually set equal to 2 and z_0 is related to the mean implantation depth given by

$$\bar{z}\ [\text{nm}] = \frac{\sqrt{\pi}}{2} z_0 = \frac{\alpha}{\rho} E^n \tag{15}$$

Here ρ represents the density of the material in g/cm^3 and α and n are two material independent constants with values of 40 g cm^{-2} keV$^{-1.62}$ and 1.62, respectively.

The S vs E and W vs E curves for Doppler broadening data are analyzed with the aid of the VEPFIT program [49] that assumes the system studied to consist of a stack of different positron trapping layers, each layer being characterized by a thickness, an S parameter, and a positron diffusion length. The S parameter for a given energy E is then given by

$$S(E) = \sum_{i=1}^{n} f_i(E)\, S_i + f_{\text{surface}}(E)\, S_{\text{surface}} + f_{\text{bulk}}(E)\, S_{\text{bulk}} \tag{16}$$

where $f_i(E)$ is the fraction of positrons implanted at energy E which after thermalization and diffusion annihilate in the layer i, S_i denotes the corresponding value of the line shape parameter in layer i, and *bulk* and *surface* refer to bulk and surface contributions. Owing to the property of linearity of the S and W parameters, one can obtain a direct interpretation of the experimental results by plotting S and W data as a trajectory in an S–W diagram. In this case the running parameter is the positron implantation energy. The advantage of the S–W diagram is that it allows one to distinguish annihilations from different types of defects, which on the basis of the S and W parameter alone could not be detected. In a Doppler broadening measurement the sensitivity to annihilations with core electrons is reduced due to the relatively high background in the spectrum. However, this background can be reduced by placing a second Ge detector opposite the first one at the same distance from the sample [50]. In this way the peak to background ratio is improved by a factor of almost 1000 compared to a single detector system at the expense of much lower count rates.

2.2.3. Positron Annihilation Lifetime

In comparison with the previously discussed positron annihilation spectroscopies, which measure projections of the momentum density of positron–electron pairs, the positron lifetime technique measures the time difference between the emission of a positron and its annihilation with an electron of the material under study. In this experiment a positron source is used which by emission of a simultaneous γ ray provides a zero time signal at the moment at which the positron enters the sample. In conventional experiments ^{22}Na emits a 1.28 MeV γ ray simultaneously with the positron, which is used to trigger the lifetime measurement. In a beam with one or more moderation stages, accelerating and retarding sections, and particle transport over an often considerable distance this is, however, not a feasible option. Several methods have been developed to overcome this problem. Early experiments used the emission of secondary electrons released from the target surface by the entrance of the positron and detected with the aid of a microchannel plate. They achieved a time resolution of ~600 ps full width at half maximum (FWHM) [51]. The resolution is limited by positrons scattering off the sample surface and returning to it after some time and by the spread in response time of the detector. A different approach consists of "bunching" the beam [i.e., collecting the positrons in the beam in groups ("bunches") equally separated in time] and subjecting these groups to fields designed to achieve a degree of time focusing. Initial experiments along these lines produced resolutions of several ns FWHM [52, 53]. Further developments in Munich [54, 55] and Japan [56] resulted in pulse widths of 150 ps with pulse repetition rates of 10^5 Hz and a time resolution of ~200 ps which is comparable to that of good conventional lifetime circuitry. Also here backscattered positrons may interfere with the lifetime measurement, either by annihilations in the surrounding materials or by causing delayed start signals.

Typical values of positron lifetimes are 100–200 ps for bulk materials, corresponding to free-positron annihilation. When the size of the open volume is large enough to accommodate a Ps atom, the lifetime is considerably higher (~ns). The lifetime spectrum can be described as a sum of decaying exponentials,

$$F(t) = \sum_{i=1}^{N} \frac{I_i}{\tau_i} \exp(-t/\tau_i) \tag{17}$$

where N represents the number of different annihilation states, τ_i the lifetime of positrons corresponding to the state i, and I_i the intensity of that state. Contemporary positron lifetime setups commonly use nonmoderated ^{22}Na sources. Therefore the information extracted from a lifetime spectrum is only globally due to the uncertainty in depth at which the positrons annihilate. However, positron lifetime equipment combined with a slow positron beam is now operational or under construction in a few laboratories [57, 58].

2.2.4. Modeling of Positron Diffusion

In experiments where the energy of the implanted positrons can be varied (beam experiments) it is important to know the depth dependent concentration profile of the thermalized positrons. From the concentration one can calculate the number of positrons per unit of time that (i) is captured in defects and subsequently annihilates, (ii) annihilates in the bulk, and (iii) arrives at the surface. To this purpose the time-averaged positron concentration is derived from the balance equation

$$D\frac{d^2c}{dz^2} - \frac{d}{dz}(v_d c) + I(z) - \textstyle\sum_j k_j n_j c - \lambda_b c = 0 \tag{18}$$

where $c(z)$ is the depth dependent concentration, $v_d(z) = \mu E(z)$ is the drift velocity with μ the positron mobility and E the electrical field strength in the material (in particular for activated semiconductor devices), $I(z)$ is the positron implantation rate at depth z, $n_j(z)$ is the density of defects of type j, k_j is the specific capture rate for capture at defects of type j, $\lambda_b = \tau_b^{-1}$ is the annihilation rate in the bulk (τ_b is the bulk lifetime), and D is the positron diffusivity.

The various terms in Eq. (18) describe the following processes: *diffusion, drift, implantation, capture*, and *annihilation*

of the positron. It should be noted that this balance equation is equally valid for other mobile particles implanted in materials. The equation may also be used for hydrogen and helium in metals in which case the electric field drift term and the annihilation term should be omitted.

When there is a uniform concentration of defects n_t with specific trapping rate κ_t and the surface is located at a sufficient distance from the depth of implantation z_f, the solution will read

$$c(z) = c(z_f)\exp[-(z - z_f)/L_+] \tag{19}$$

where $L_+ = [D^+/(\kappa_t n_t + \lambda_b)]^{1/2}$ is the effective diffusion length (i.e., the averaged distance traveled by the positron before it is trapped or annihilated). From Eq. (19) the boundary condition at a large distance from the surface reads

$$\left[\frac{dc}{dz}\right]_{z=z_f} = -c(z_f)/L_+ \tag{20}$$

For the outer surface a similar boundary condition can be defined, but now the absorption length L_a instead of the effective diffusion length L_+ is involved

$$\left[\frac{dc}{dz}\right]_{z=0} = -c(z=0)/L_a \tag{21}$$

The stationary diffusion equation can be solved numerically and the solutions constitute the basis for the analysis of positron annihilation measurements with variable energy positron beams.

2.3. Intense Positron Sources

Positron spectroscopies for depth resolved material analysis, at the surface or near surface, require monochromatic positron beams. Low intensity ($<10^6$ e^+/s) positron beams based on commonly used radioactive isotope sources used in positron physics (^{22}Na, ^{58}Co) have been extensively applied to the investigation of surface or near surface material properties (e.g., surface structure or depth distribution of defects at surfaces and buried interfaces). Low-intensity beams are sufficient to perform many of the positron annihilation spectroscopies (e.g., Doppler broadening of annihilation radiation and positron lifetime), but the application of the 2D-ACAR method and positron microscopy to the study of surfaces and layered systems had to wait for the development of high-intensity slow positron beams ($\sim 10^8$ e^+/s). Nowadays, such beams in which positrons are created by pair production induced by high-energy γ rays are operational or under construction in several laboratories [8], and as these facilities are commissioned the number of applications of positron annihilation spectroscopies in the field of materials science will increase.

A slow positron beam system usually consists of the following components: (i) a primary positron source with a continuous energy distribution, (ii) a positron moderator located near the primary source, and (iii) a beam transport system. In order to increase the intensity of a slow positron beam one or more of these components will have to be improved. The beam transport systems are highly efficient so there is little room for improvement. The role of the positron moderator is to slow down the fast positrons [e.g., when fast positrons are implanted into a well-annealed sample, those within the diffusion length ($\sim$100 nm) from the surface can diffuse back and are ejected nearly monoenergetically as a result of their negative work function]. Typical moderating efficiencies are modest, of the order of $\sim 10^{-4}$–10^{-3}, and these limitations are based on the physics of positron interaction in the moderating system and the geometry of the moderators. The major factor to achieve intense slow positron beams is therefore the primary source strength, which has practically no physical limit. In the following, different approaches adopted at different laboratories to achieve strong sources of moderated positrons are presented.

2.3.1. Sources Based on Relatively Short-Lived Radioactive Isotopes

Reactor Based Sources A high-current positron beam can be achieved by activation of ^{63}Cu in a nuclear reactor to ^{64}Cu, which decays under the emission of a positron. Lynn et al. [59] built an intense beam relying on the activation of a strong (800 Ci) source of copper in the High Flux Beam Reactor at Brookhaven. The positron emitting ^{64}Cu, obtained via the nuclear reaction ^{63}Cu$(n, \gamma)^{64}$Cu, was evaporated onto a single-crystal W(110) substrate in an external vacuum system to provide a self-moderating copper film. The system has operated successfully for several years and has produced many useful data for atomic and solid-state physics including 2D-ACAR on thin films. Disadvantages were the operational efforts. Every three days the source had to be refreshed due to the 12.8 h decay time of the ^{64}Cu radioisotope.

Accelerator Based Sources Irradiation of materials with accelerator ion beams can produce intense positron emitting radionuclides [60]. An example is the irradiation of ^{12}C with a high current of low-energy deuterium ions, which produces ^{13}N. By interchanging the carbon target between the deuteron beam and a positron moderation stage an intense quasi-continuous slow positron beam can be obtained. Since compact linear accelerators are available, this technique is promising for producing inexpensive high-current positron beams. A first attempt to produce such a beam was undertaken at Brandeis University [8].

2.3.2. Sources Based on Pair Production

Pair production is the most commonly used method for producing high-current positron beams. When photons with energy higher than 1.022 MeV interact with the fields surrounding heavy nuclei (e.g., tungsten) that energy can be converted into e^+–e^- pairs. The probability of pair production increases when the photon energy increases. The fast positrons are then separated from the electrons and moderated in order to obtain an intense slow positron beam. Two primary designs are available: (a) linac based positron beams and (b) nuclear reactor based positron beams.

Linac Based Positron Beams Successful intense positron beams are linear accelerator (linac) based positron beams with intensities over 10^9 e^+/s. Positrons are generated in a high-Z target (e.g., tungsten) hit by high-energy (~100 MeV) electrons. Bremsstrahlung photon generation and subsequent pair formation yield fast positrons that can be moderated by tungsten foils. An example of such an intense positron beam was developed at the 100 MeV electron linac at Lawrence Livermore National Laboratory [61]. Other designs are available or under development at different laboratories [7]. Linac based positron beams are inherently pulsed. A quasi-continuous beam can be obtained by trapping the positrons from the linac in a magnetic bottle and then releasing them over the period between linac beam pulses.

Reactor Based Positron Beams The advantage of these beams is that they are practically continuous over time. Pair production can be obtained in nuclear reactors in two different ways: (i) *capture of high-energy gamma rays in thin tungsten moderator foil*s and (ii) *neutron capture of gamma rays from cadmium foils*. The latter is an alternative source of high-energy photons for pair production of positrons. It consists of the gamma rays generated when thermal neutrons are captured in cadmium foils via the nuclear reaction ^{113}Cd (n, γ) ^{114}Cd. Consequently, the gamma rays produce electron–positron pairs in tungsten foils. An intense beam facility based on neutron capture gamma rays is under development at the Research Reactor Munich [62].

2.4. Transmission Electron Microscopy

Transmission electron microscopy is a valuable tool to study defects because it allows for the observation and analysis of materials down to the (sub)nanometer scale. The resolution of a light microscope is limited to the order of a micrometer because of the wavelength of light; an electron microscope is limited by its lenses and stability and may reveal details down to the atomistic scale. The contrast is formed mainly by the elastic scattering of electrons with matter and can be used to obtain real-space as well as diffraction information. The inelastic scattering of some of the electrons can be helpful in obtaining chemical information from very small regions in the specimen. Magnetic lenses are used to focus the electrons and to a large extent the aberrations of these lenses still limit the resolution of the electron microscope. By reduction of the spherical aberration, an increase of the mechanical stability and the use of higher voltages (~400 kV) led to the development of the dedicated high resolution transmission electron microscope (HRTEM). With this type of microscope it is possible to observe the atomistic structure of crystalline solids in real space.

In a HRTEM a crystalline material under investigation is thinned to a thickness of a few nanometers in such a way that it becomes transparent for high-energy electrons. Electrons are then accelerated to typically 400 kV and focused on the specimen. The exiting electron wavefront is magnified and projected on a phosphor screen or a recording medium, which results in a projected 2D image of the atomic structure of the material. The technique of HRTEM originated in the technique of phase contrast microscopy, introduced by Zernike [63] for optical microscopy at the University of Groningen. In 1953 he received the Nobel Prize in physics for this invention. High-resolution TEM imaging is based on the same principles. Phase-contrast imaging derives contrast from the phase differences among the different beams scattered by the specimen, causing addition and subtraction of amplitude from the forward-scattered beam. Components of the phase difference come from both the scattering processes itself and the electron optics of the microscope [64]. Electron microscopy has been around for a number of decades but with ongoing development of the instrument the point-to-point resolution has now dropped well below 0.2 nm, making it possible to distinguish the atomic structure of many materials along low-index crystallographic directions.

A transmission electron microscope consists of an illumination system, specimen stage, and imaging system, analogous to a conventional light microscope. In this research two different electron microscopes are used, a JEOL 4000 EX/II (LaB_6, 400 kV) and a JEOL 2010F (FEG, 200 kV). Several textbooks and reviews have been written on the subject of (high-resolution) microscopy, sadly with different notation conventions. This text adopts the notation used in [64, 65]. Spence [66] has written a more in-depth description of imaging in HRTEM theory. Williams and Carter [67] have written a very clear and complete textbook on general aspects of electron microscopy.

The conventional way to generate electrons is to use thermionic emission. Any material that is heated to a high enough temperature will emit electrons when they have enough energy to overcome the work function. In practice this can only be done with high melting materials (such as tungsten) or low work function materials like LaB_6. Richardson's law describes thermionic emission as a function of the work function Φ and temperature T,

$$J = AT^2 e^{\frac{\Phi}{kT}} \tag{22}$$

with J the current density at the tip and A the so-called Richardson's constant, depending on the material used. In electron microscopes tungsten filaments were most commonly used until the introduction of LaB_6. These LaB_6 sources have the advantage of a lower operating temperature, which reduces the energy spread of the electrons and increases the brightness.

Another way of extracting electrons from a material is to apply a high electric field to the emitter that enables the electrons to tunnel through the barrier. Sharpening the tip may enhance the electric field because the electric field at the apex of the tip is inversely proportional to the radius of the apex,

$$E = \frac{V}{kR} \tag{23}$$

with k a correction factor for the tip geometry (usually $\approx$2). The advantage of this cold field emitter gun (cold FEG) is that the emission process can be done at room temperature, reducing the energy spread of the electrons. The small size of the emitting area and the shape of the electric field results in a very small (virtual) source size in the order of

nanometers with brightness that is three orders of magnitude higher than for thermionic sources. It is thus possible to focus the beam to a very small probe for chemical analysis at an atomistic level or to fan the beam to produce a beam with high spatial coherence over a large area of the specimen. The disadvantages of this type of emitter are the need for (expensive) ultrahigh vacuum equipment to keep the surface clean, the need for extra magnetic shielding around the emitter, and a limited lifetime.

Some of these disadvantages of cold FEG emitters can be circumvented by heating the emitter to moderate temperatures (1500 °C) in the case of a thermally assisted FEG or by coating the tip with ZrO_2, which reduces the work function at elevated temperatures and keeps the emission stable (Schottky emitter). This causes a doubling in the energy spread of the electrons and some reduction in brightness. The Schottky emitter is most widely used in commercial FEG-TEMs because of its stability, lifetime, and high intensity.

After leaving the electron gun the electrons are accelerated toward the anode and enter the column. Because the velocity of the electrons approaches the speed of light the wavelength of the electrons has to be corrected for relativistic effects

$$\lambda = \left(\frac{2m_0 eE}{h^2} \left(1 + \frac{eE}{2m_0 c^2} \right) \right)^{1/2} \tag{24}$$

For 400 kV electrons $\lambda = 1.64$ pm, which is much smaller than the resolution of any electron microscope because the resolution is limited by aberrations of the objective lens and not by the wavelength of the electrons. The electron beam is now focused on the specimen with the condenser lenses and aligned using several alignment coils. The function of the condenser lens system is to provide a parallel beam of electrons at the specimen surface. In practice this is not possible and the beam always possesses a certain kind of convergence when imaging at high resolution, usually in the range of 1 mrad for LaB_6 emitters and 0.1 mrad for FEGs.

After entering the specimen most of the electrons are elastically scattered by the nuclei of the atoms in the specimen. Some electrons are inelastically scattered by the electrons in the specimen. Compared to X-ray or neutron diffraction, the interaction of electrons with the specimen is huge and multiple scattering events are common. For thick specimens at lower resolutions an incoherent particle model can describe the interaction of the electrons with the specimen. However, with thin specimens at high resolution this description fails because the wave character of the electrons is then predominant. The electrons passing the specimen near the nuclei are somewhat accelerated toward the nuclei, causing small, local, reductions in wavelength, and this, in turn, results in a small phase change of the electrons. Information about the specimen structure is therefore transferred to the phase of the electrons.

Only elastically scattered electrons are of importance in formation of high-resolution images. The inelastic scattered electrons contribute mostly to the background of the image. These electrons can be removed by inserting an energy filter in the microscope between the specimen and recording device. Only the elastically scattered electrons are now used for the image formation. The energy loss spectrum of the inelastic scattered electrons contains valuable information about the chemical composition of the specimen. This information can be extracted by observing the electron energy loss spectrum or by filtering parts of the loss spectrum. The inelastic scattered electrons also produce Kikuchi lines in the electron diffraction pattern that are helpful for accurate crystallographic alignment of the crystals in the specimen.

To retrieve structural information on the specimen from the micrographs it is necessary to calculate the trajectory of the electron wave through the specimen. In kinematic approximation, multiple scattering of the electrons in the sample is ignored and this results in an undisturbed central beam. This approach already fails at a small thickness or a single atom. In dynamical calculations all the scattered beams and their mutual exchange of intensity during the course of multiple scattering in the specimen are taken into account. It is possible to do full dynamical calculations but these are soon limited by the available computing power. Using the fact that the vast majority of the electrons are scattered in a forward direction with small diffraction angles, Cowley and Moodie [68] devised the multislice approximation. Here the 3D crystal is divided into multiple thin slices perpendicular to the electron beam; in between the electron wave propagates. The incident electron wavefront intensity is described by a plane wave with wave vector k propagating *in vacuo*,

$$\begin{aligned} \Psi_0(\mathbf{r}) &= e^{-2\pi i \mathbf{kr}} \\ k = |\mathbf{k}| &= \frac{\sqrt{2meE}}{h} = \frac{1}{\lambda} \end{aligned} \tag{25}$$

with λ the nonrelativistic wavelength of the electrons. The refractive index in every slice depends on the position in the slice because the electrons will be accelerated toward the nuclei, whereby their wavelength is reduced. This phase change can be expressed by multiplying $\Psi_n(\mathbf{r})$ with a transmission function $q(x, y)$, assuming that the specimen acts as a pure phase object,

$$q(x, y) = e^{-i\sigma\phi(x, y)} \tag{26}$$

with $\phi(x, y)$ the projected potential distribution in the slice, averaged over the electron beam direction and $\sigma = \pi/\lambda E$ the interaction constant. Absorption is neglected in this approach, the so-called phase-object approximation. A further simplification can be made by only considering the first-order approximation of Eq. (26)

$$q(x, y) = 1 - i\sigma\phi(x, y) \tag{27}$$

This is known as the weak phase-object approximation in which the amplitude information is linearly related to the projected potential. The propagation between the slices, each of thickness δ, can be described by convoluting (26) with the Fresnel propagator,

$$p(x, y) = e^{-\pi i(x^2+y^2)/\lambda\delta} \tag{28}$$

and the emerging spherical wavefront is approximated by a paraboloidal wavefront.

The nth wave function can now be obtained by a convolution,

$$\Psi_n(x, y) = (\Psi_{n-1} \otimes p_{n-1}(x, y)) \cdot q_n(x, y) \tag{29}$$

with p_{n-1} the propagator from slice $n-1$ to slice n. In this way the exit wave computed can subsequently be used as input for the image formation calculations. Usually the convolution is done by a multiplication in Fourier space, which saves computing time. After the exit wave has left the specimen the electron wavefront has to be converted into an image. Because the exit wave mainly contains phase information, these phase differences have to be converted to intensities, much like the phase contrast microscope invented by Zernike [63]. The exit wave can be described in reciprocal space by performing a Fourier transform. The effect of the spherical aberration of the objective lens is that the electrons entering the objective lens at different distances from the optical axis are focused at different planes and also travel different distances, causing a phase difference for the different scattered electrons. The phase factor $\chi(U)$ describes the phase difference introduced by the defocus and spherical aberration,

$$\chi(U) = \pi\lambda\Delta f U^2 + \frac{1}{2}\pi C_s\lambda^3 U^4 \tag{30}$$

with Δf the defocus value of the objective lens, U the distance of the reciprocal lattice point from the optic axis, and C_s the spherical aberration. In reciprocal space this phase factor can be accounted for by multiplying the exit wave with $B(U)$, which is equivalent to a convolution in real space:

$$B(U) = \exp(i\chi(U)) \tag{31}$$

If the specimen behaves as a weak phase object only the imaginary part of $B(U)$ contributes to the intensity of the image, $T(U)$,

$$\begin{aligned} T(U) &= 2\,\sin(\chi(U)) \\ &= 2\,\sin\left(\pi\lambda\Delta f U^2 + \frac{1}{2}\pi C_s\lambda^3 U^4\right) \end{aligned} \tag{32}$$

which is usually referred to as the objective lens transfer function if the specimen behaves like a weak phase object. Only the real part is considered because the phase information from the specimen is converted in intensity information by the phase shift of the objective lens. In the microscope an aperture is inserted in the back focal plane of the objective lens that transmits beams only to a certain angle. This can be represented by an aperture function $A(U)$ which is unity for $U < U_0$ and zero outside this radius. When $T(U)$ is negative, the atoms in the specimen would appear as dark spots against a bright background and vice versa. For $T(U) = 0$ no contrast results. An ideal behavior of $T(U)$ would be zero at $U = 0$ (very long distances in the specimen) and $U > U_0$ (frequencies beyond the aperture size) and large and negative for $0 < U < U_0$.

The contrast of a high-resolution image depends strongly on the microscope settings and parameters. In practice, not all the information in $T(U)$ is visible in the image. This is caused by electrical instabilities in the microscope creating a spread of focus because of the chromatic aberration of the objective lens, which results in damped higher frequencies. Mechanical instabilities and energy loss due to inelastic scattering of the electrons by the specimen also contribute to the spread in defocus. The inelastic scattered electrons contributing to the image can be removed by inserting an energy filter in the microscope between the objective lens and the image recording media. Another factor that damps the higher frequencies is the beam convergence. Because the electron beam has to be focused on a small spot on the specimen there is some convergence of the beam present. This also affects the resolution because the specimen is now illuminated from different angles at the same time. These effects that affect the resolution can be represented by multiplying $\sin(\chi)$ by the damping envelopes E_α and E_Δ, which represent the damping by the convergence and spread in defocus respectively:

$$\begin{aligned} E_\Delta(U) &= \exp\left[-\frac{1}{2}\pi^2\lambda^2\Delta^2 U^4\right] \\ E_\alpha(U) &= \exp[-\pi^2\alpha^2(\varepsilon + \lambda^2 C_s U^2)^2 U^2] \end{aligned} \tag{33}$$

The resulting contrast transfer function (CTF) for the JEOL 4000 EX/II is plotted in Figure 6 with the damping envelops E_α and E_Δ. For higher frequencies the CTF is now damped and approaching zero. It becomes clear that defining a resolution for the HRTEM is not obvious. Several different resolutions can be defined as stated by O'Keefe [69]:

(1) Fringe or lattice resolution: This is related to the highest spatial frequency present in the image. In thicker crystals second-order or nonlinear interference may cause this. As the sign of $\sin(\chi)$ is not known there is in general no correspondence between the structure and the image. This resolution is limited by beam convergence and spread in defocus.
(2) Information limit resolution: This is related to the highest spatial frequency that is transferred linearly to the intensity spectrum. These frequencies may fall in a passband, with blocked lower frequencies. Usually this resolution is almost equal to the lattice resolution.
(3) Scherzer resolution or point resolution: This is the highest possible resolution when no lower frequencies are blocked or passed with opposite sign. This is the most relevant resolution definition because it results in a projected image of the specimen structure. The exact definition for the Scherzer defocus varies slightly in literature. In general, for images taken at optimum defocus of $(\frac{4}{3}C_s\lambda)^{1/2}$ the highest transferred frequency is equal to $1.51C_s^{-1/4}\lambda^{-3/4}[\mathrm{m}^{-1}]$.

In Figure 6 the CTFs of the JEOL 4000 EX/II, JEOL ARM 1250, and JEOL 2010F at Scherzer defocus are plotted with the damping envelopes. The CTF of the 2010F shows the rapid oscillations of the FEG due to higher spatial coherence. In this case there is a big gap between point resolution and information limit. For the ARM 1250 microscope at the Max Planck Institute für Metallforschung in Stuttgart the damping envelope is limited by the spread-of-focus and the point resolution is almost equal to the information limit. Using a FEG will not substantially increase the resolution here. Generally speaking, using a FEG in

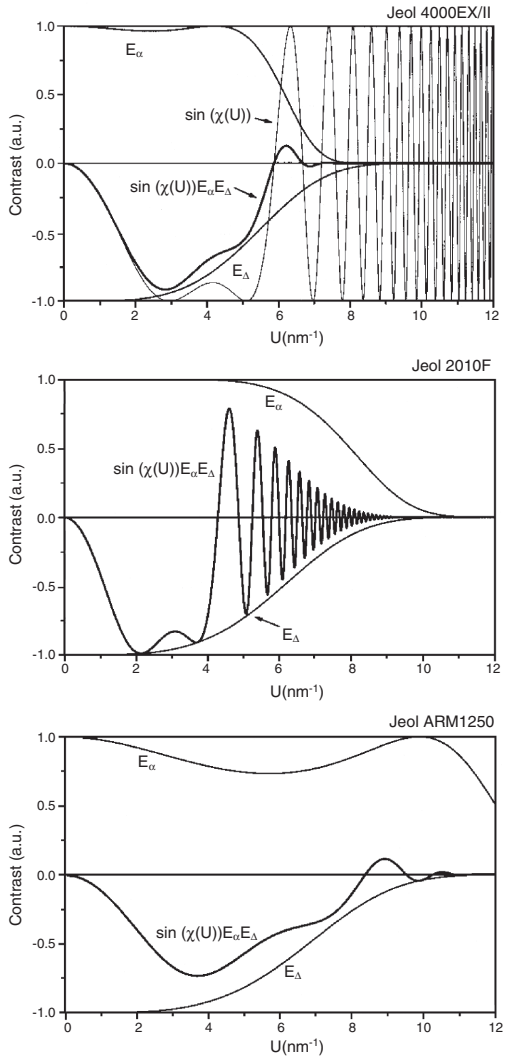


Figure 6. Contrast transfer functions (thick lines) and damping envelopes of the used microscopes at Scherzer defocus of 47, 58, and 51 nm respectively. Note the many oscillations in the CTF of the 2010F.

lower voltage ($\leq$200 kV) instruments helps to increase the information limit because in these instruments beam convergence dominates the damping envelope. With higher voltage instruments the higher acceleration voltage increases the brightness of the source and the damping envelope is limited by the spread of defocus. The spherical aberration of the objective lens can be lowered but generally at the expense of a decrease in tilt capabilities of the specimen. It is possible to compensate C_s by a set of hexapole lenses as suggested by Scherzer [70] in 1947 but it became feasible only recently due to the complex technology involved. Haider et al. demonstrated a C_s corrected 200 kV FEG microscope [71] in which C_s was set at 0.05 mm to reach an optimum between contrast and resolution. The point resolution of this microscope was now equal to the information limit of 0.14 nm. Successful application of this technology would put the limiting factors of the microscope at the chromatic aberration and mechanical vibrations that currently limit the resolution around 0.1 nm.

For correct interpretation of the structure the specimen has to be carefully aligned along a zone axis, which is done with the help of Kikuchi patterns and an even distribution of the spot intensities in diffraction. The misalignment of the crystal is in this way reduced to a fraction of a milliradian. Beam tilt has a more severe influence on the image because the tilted beam enters the objective lens at an angle that causes phase changes in the electron wavefront. To correct the beam tilt the voltage center alignment causes the acceleration voltage to be varied in order to find the center of magnification that is then placed on-axis. This alignment might not be correct due to misalignments in the imaging system after the objective lens. For better correction of the beam tilt the coma-free alignment procedure is necessary. This is done by varying the beam tilt between two values in both directions and optimizing the values until images with opposite beam tilts are similar. This procedure is only practically feasible with a computer-controlled microscope equipped with a charge coupled device camera and is not commonly used; instead, the voltage center alignment used could cause some residual beam tilt.

3. DEFECTS AND NANOCLUSTER ENGINEERING IN MgO

3.1. Introduction

Currently much research in the field of defect engineering is devoted to defects in magnesium oxide. Applications of MgO include the development of optical communications media [72] and the use of MgO in nuclear reactor fuels. An absorption band centered around 550 nm, known as Mie resonance, has been observed after Au/Ag implantation in a number of studies [73, 74]. Copper, silver, and gold are the most common metals used for ion implantation. The formation of metallic nanoprecipitates in MgO after implantation and subsequent annealing has been investigated for various annealing treatments using techniques such as Rutherford Backscattering Spectrometry/Channeling (RBS-C) and optical absorption spectrometry [75]. A related topic of research is the role of cavities in the material, which may act as gettering centers for implanted metal ions. Recently nanovoids in MgO have been created by helium implantation and subsequent annealing [76].

Here, we discuss investigations about the nucleation, aggregation, and dissociation of metallic nanoprecipitates and their possible interaction with vacancy-type defects. For this purpose, three complementary techniques were used. Positron beam analysis (PBA) is a very sensitive tool to monitor depth resolved vacancy-type defects [77]. RBS can provide information about the depth distribution of the implanted copper ions. Finally cross-sectional transmission electron microscopy (X-TEM) can provide information about the configuration of the metallic nanoprecipitates and vacancy-type defects.

Conventionally metal nanoclusters are formed after 1–2 MeV implantation followed by high-temperature anneals. An alternative way to form nanoclusters was studied in [78]. Prior to the metal ion implantation a chain of nanosize cavities was created as a result of 30 keV He^+

ion implantation followed by annealing. In the complementary TEM study [79] it was observed that the cavities are elongated in the [100] direction and have a rectangular form. Subsequently, Au ions were introduced by low-energy, 30 keV implantation. Then the samples were annealed to provide diffusion of the metal ions to the cavity layer. However, only a small fraction of Au reached the cavities located at a depth of about 200 nm. Most of the Au ions created a band of clusters close to the surface within the implantation range. In the present study MgO samples containing a band of nanocavities have been implanted with high-energy Au and Ag ions. The energies were chosen to introduce the metal ions directly in the vicinity of the cavity band. Single crystal MgO (100) samples with dimensions $10 \times 10 \times 1$ mm^3 were used. The samples were treated in the following way: The nanosize cavities were formed by 30 keV $^3He^+$ implantation followed by annealing in air at 1273 K for 0.5 h. After this the samples were implanted with 30 keV Au at room temperature (RT) and 1273 K. Another set of four samples with cavities was implanted with 600 keV Ag and 1 MeV Au ions at RT and at 873 K. After the implantation the samples were stepwise annealed in air in order to form the nanoclusters. The optical properties of the samples were studied by light absorption. The structure of the defects introduced by implantation was studied by X-TEM and PBA. A detailed description of the PBA technique is presented in Section 2 [51]. In short, with PBA the sample is probed by a monoenergetic beam of positrons with energies ranging from 0.1 to 25 keV, which correspond to typical depths from the surface up to 3000 nm. Annihilation of a positron with an electron at rest results in two photons of 511 keV each. If the electron has a nonzero momentum the energy of the photons is Doppler shifted resulting in a broader energy distribution. The energy spectrum of the 511 keV photon annihilation peak is characterized by two parameters: S and W (see also Section 2). The S parameter is defined as the contribution of the central part of the peak and is related to annihilations with valence or conduction electrons. The W parameter is defined by the contribution of the "wings" in the spectrum and is related to annihilations with high momentum core electrons. Thus, high S values are usually attributed to vacancy related defects where the chance for a positron to annihilate with a high-momentum core electron is low. Another contribution is positronium, which can be formed in cavities of nanometer size.

3.2. Results and Discussion

The bands of nanosize cavities and Au nanoclusters observed in X-TEM are shown in Figure 7. In this work an MgO sample containing a band of cavities at a depth of about 200 nm was implanted with 30 keV Au ions to a dose of 10^{16} cm^{-2}. The sample was subsequently annealed in air at 1273 K. It is seen that a band of nanoclusters is formed within 25–75 nm of the surface. The cavity layer formed prior to the Au implantation had developed into a band of elongated internal cracks during the annealing treatment. The cracks are presumably formed by merging a few adjacent cavities into one. Additionally, the Au implantation has created another band of rectangular nanosize cavities within the first 25 nm. The fact that the implantation

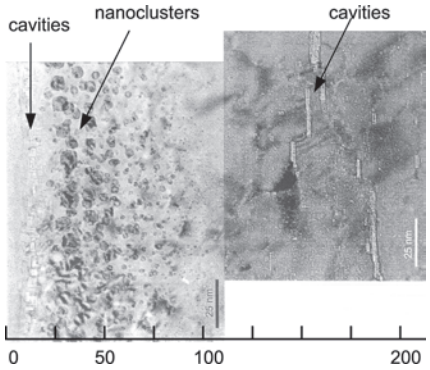


Figure 7. X-TEM picture of two bands of cavities and a band of nanoclusters obtained after a combination of 30 keV He and Au implantations and anneals discussed in the text. The depth location of the bands is shown in nanometers.

induced vacancies have agglomerated into cavities rather than escaping at the surface suggests a very short diffusion length within the damage region. The Au nanoclusters were later studied by X-ray diffraction (XRD) and X-TEM and were found to be semicoherent with the MgO host lattice [80].

In order to introduce the metal ions directly into the vicinity of the cavity band the implantations were carried out at energies which correspond to the depth of the original cavity band (i.e., 200 nm). Thus, the implantation energies for Au and Ag ions were 1 MeV and 600 keV, respectively. To enhance diffusion two samples were implanted at an elevated temperature of 873 K. The light absorption measurements performed directly after Au and Ag implantations are shown in Figures 8 and 9, respectively. For as-implanted samples an absorption peak at 560 nm is observed only in the case of RT implantation (curve a in Figure 8 and curve a in Figure 9). This absorption band has disappeared after the

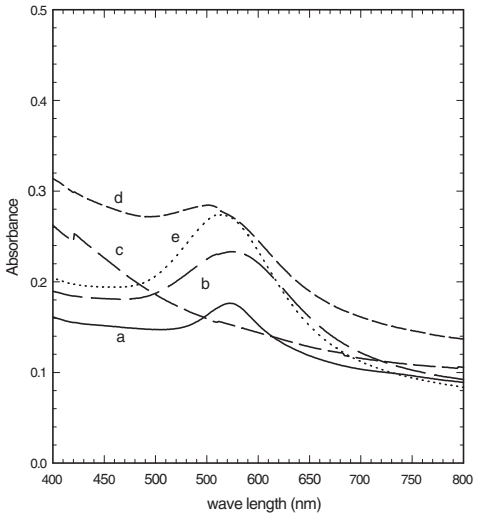


Figure 8. Light absorption measurements of the MgO samples after 1000 keV, 10^{16} cm^{-2} Au ion implantation: (a) as-implanted at RT; (b) implanted at RT and annealed at 1473 K, 23 h; (c) as-implanted at 873 K; (d) implanted at 873 K and annealed at 1473 K, 23 h; (e) implanted at RT and annealed at 1473 K, 22 h (no cavities).

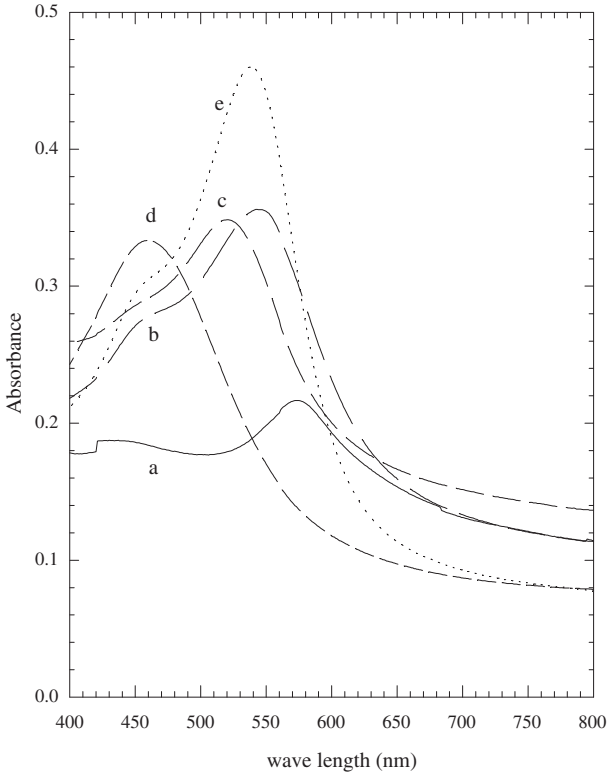


Figure 9. Light absorption measurements of the MgO samples after 600 keV, 10^{16} cm^{-2} Ag ion implantation: (a) as-implanted at RT; (b) implanted at RT and annealed at 1473 K, 23 h; (c) implanted at 873 K and annealed at 1473 K, 23 h; (d) implanted at RT and annealed at 1273 K, 5 h; (e) implanted at RT and annealed at 1473 K, 22 h [(d) and (e) contain no cavities].

samples were annealed at 873 K. That suggests that these absorption centers are introduced by the implantation and can be related to V-centers. During further anneals up to 1473 K nanoclusters are formed. Mie resonance peaks are observed at 570 nm for Au and at 540 nm for Ag. In the case of the Ag implantation another peak, located at 460 nm, can be observed. However, no evident enhancement of the metal nanocluster formation was observed in the samples containing nanocavities.

The evolution of the metal ion precipitation was also monitored with the help of PBA. The experimental data are presented on the *SW* map. As the positron implantation energy increases the experimental points form a path, which runs between a number of so-called cluster points. Each cluster point is related to a certain (electron) environment where positron annihilation takes place Thus, in Figure 10 four cluster points are identified (shown with open circles). For the case of the cavity containing sample (solid circles), the path runs first from the "surface" to the "MgO bulk"-cluster point. As the energy increases the positrons are implanted in the cavity layer and the *SW* path is deviated to the cluster point denoted as "cavities." At higher positron energies the data points tend to cluster around the "MgO-bulk" point again. After the sample was implanted with 1000 keV Au ions and annealed another cluster point ascribed to Au nanoclusters is observed. For the case of the Ag implanted sample (shown in Fig. 11) the

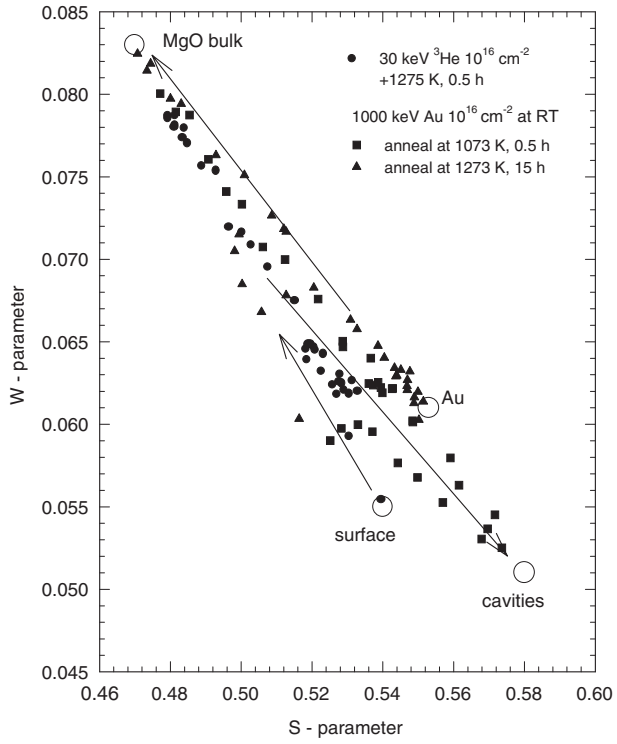


Figure 10. *SW* map of MgO with cavities implanted with 1000 keV Au and subsequently annealed at 1073 and 1273 K. Cluster points shown on the map with opened circles are related to positrons trapped at the surface, cavities, nanoclusters, and defect-free MgO bulk.

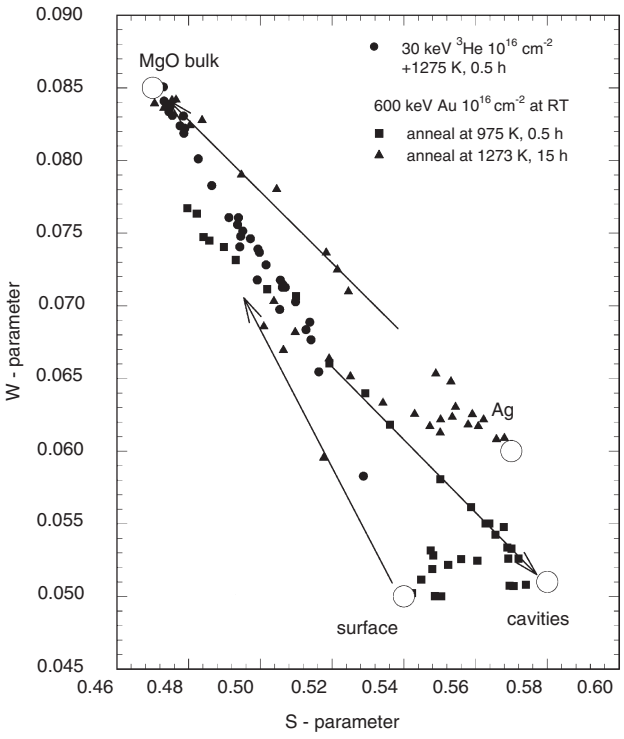


Figure 11. *SW* map of MgO with cavities implanted with 600 keV Ag and subsequently annealed at 973 and 1273 K. Cluster points shown on the map with opened circles are related to positrons trapped at the surface, cavities, nanoclusters, and defect-free MgO bulk.

cluster point ascribed to Ag nanoclusters is identified. Note that while the "cavities" and "MgO-bulk" cluster points in both cases are the same; the cluster points related to Au and Ag nanoclusters are different.

In order to obtain the depth location of the defected layers the *S* curves measured prior and after the metal ion implantations were analyzed with the program VEPFIT [49]. The *S* curve (a) ascribed to the original band of cavities was fitted with a model containing 3 layers: a layer of cavities between two layers of defect-free MgO matrix. The *S* curve (b) measured after the Au implantation was fitted with a five layer model: three defected layers (two with cavities and one with metal precipitates) and two defect-free layers. The results of the fit are shown in Figure 12. The stopping ranges of ions obtained by the TRIM [81] simulations are also shown for comparison. The implantation energies of He, Au, and Ag ions were chosen to yield the same implantation range. However, a discrepancy within 100 nm is observed between the predicted by TRIM ranges and the location of the defected layers obtained by VEPFIT. From the PBA studies alone, however, is difficult to say whether the metal ions have precipitated into a cluster or are still dissolved in the matrix. It can be concluded that nanosize cavities developed after He implantation were used as precursors for the precipitation of the metal nanoclusters. Metal ions were introduced by low-energy (30 keV) Au implantation and high-energy Au (1 MeV) and Ag (600 keV) implantations. In the case of low-energy implantation nanoclusters are predominantly formed within the implantation range close to the surface. In the case of high-energy implantation formation of the nanoclusters was observed by the presence of Mie absorption bands. However, the location of the nanocluster layer has yet to be identified.

Nanocluster sizes can be estimated using three different methods: using the Doyle formula for the broadening of the optical absorption peak associated with Mie plasmon resonance, using the Scherrer formula for the broadening of the Au and Ag XRD peaks, and from direct observation of TEM images. The average size and the size distribution are best determined from X-TEM. However, this technique often implies destruction of the sample and the specimen preparation is time consuming. Therefore, many authors working on annealing studies try to estimate the precipitate size from the broadening of the optical absorption peak that is present as a result of Mie plasmon resonance [82]. Here the Doyle theory [83] is used to estimate the nanocluster size during the annealing procedure. Alternatively, the size can be estimated from the broadening of the XRD diffraction peaks using the Scherrer formula [84]. Here, we compare the nanocluster size as obtained from the three methods mentioned. XRD and X-TEM are also employed in order to determine the coherency and the orientation relationship of the gold clusters with the surrounding matrix. The sample treatment data are listed in Table 1. In the Doyle theory, the free mean path of the electrons constituting the plasmon is limited by the size of the nanocluster. This allows estimation of the cluster size as [83]

$$D = 2\frac{v_F}{\Delta\omega_{1/2}} = 2\frac{v_F}{2\pi c}\frac{\lambda_{\max}^2}{\Delta\lambda_{1/2}} \tag{34}$$

Here D is the diameter of the precipitate (m), v_F is the Fermi velocity (1.39×106 m/s both for gold and silver [85]), $\Delta\omega_{1/2}$ is the FWHM of the peak when plotted as a function of the circular frequency ω (rad/s), c is the speed of light (m/s), $\lambda_{\max}$ is the position of the centroid of the peak, and $\Delta\lambda_{1/2}$ is the FWHM of the peak when plotted as a function of wavelength (m). The background in the optical spectra was fitted with an exponential function and the Mie plasmon resonance peak was fitted with a Gaussian. The FWHM of the Gaussian and the parameters mentioned were substituted in Eq. (34), yielding cluster diameters of 4.3 and 5.1 nm for the Au and Ag clusters, respectively. High-resolution XRD measurements were carried out on the MgO(002), (113), (004), and (024) diffraction peaks.

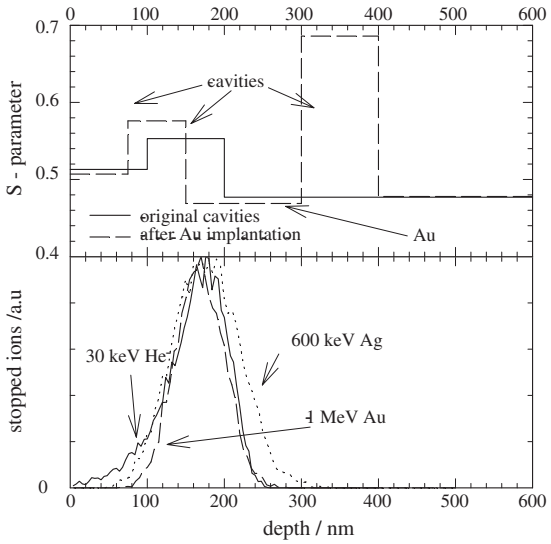


Figure 12. The results of the fitting performed with the VEPFIT program: Defect layers characterized by different *S* values are shown as a function of depth. Ranges of the implanted ions with the indicated energies are also shown in the bottom graph.

Table 1. Sample treatment and structural properties of the Au and Ag nanoclusters.

		MgO:Au	MgO:Ag
Sample treatment			
implantation		1 MeV, 10^{16} Au^+ cm^{-2}	600 keV, 10^{16} Ag^+ cm^{-2}
thermal anneal		1473K, 22 hrs	1473K, 22 hrs
Cluster size and shape			
optical abs. (Doyle)		4.3 nm	5.1 nm
XRD (Scherrer), (113) peak		4.0 nm ($K_w = 0.8290$)	9.8 nm ($K_w = 0.8863$)
XTEM	mean size	4.6 nm	11 nm
	distribution	2–14 nm	6–24 nm
	shape	spherical	octahedral
Lattice parameter *a*			
literature		4.078 Å	4.085 Å
misfit ($a_{MgO} = 4.212$ Å)		3.2%	3.0%
experimental *a* (XRD)		4.09 Å	4.12 Å
coherency strain $\varepsilon_{[100]}$ (XRD)		+0.4%	+0.9%

Figure 13a and b shows the contour plots of the (113) diffraction peaks of the MgO:Au and MgO:Ag samples. The $k_{[001]}$ axis corresponds to diffraction along the [001] crystal axis (perpendicular to the surface) and the $k_{[110]}$ axis corresponds to diffraction along the [110] crystal axis (parallel to the surface). The crystal structure of MgO, Au, and Ag is face-centered cubic (fcc) and their lattice parameters are listed in Table 1. Because the lattice parameters of Au and Ag are close to that of MgO, Au(113) and Ag(113) satellite peaks are observed near the intense MgO(113) peaks in Figure 13. The broadening of the MgO diffraction peaks is due to mosaic broadening (rocking curve). The Au and Ag satellite peaks were also found near the MgO (002), (004), and (024) diffraction peaks, clearly indicating a cube-on-cube orientation relationship with the MgO host matrix.

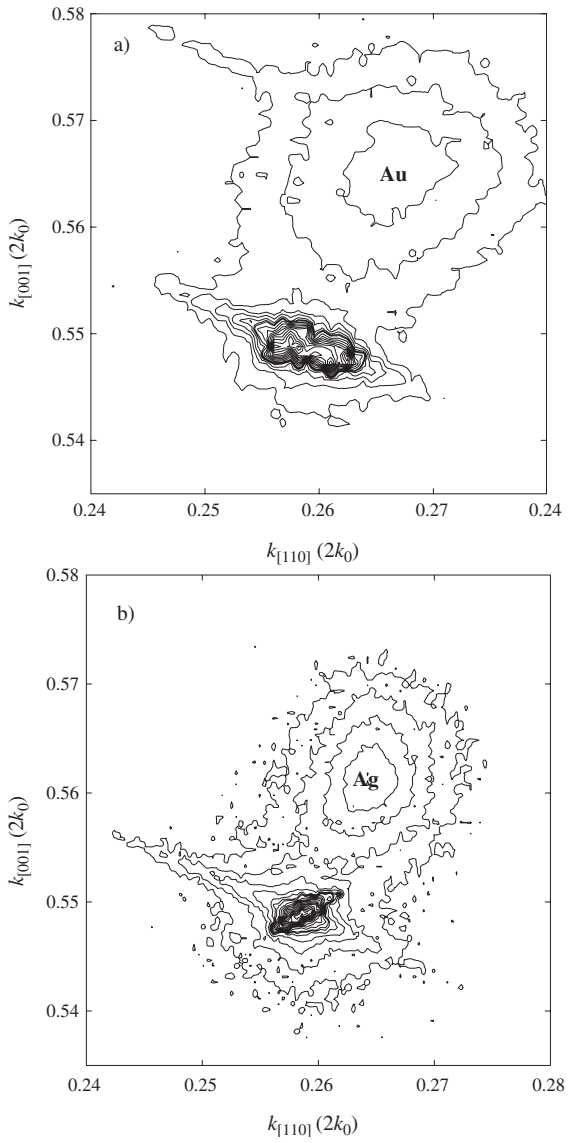


Figure 13. Reciprocal space plots of (a) the MgO(113) and Au(113) diffraction peaks of the MgO:Au sample and (b) the MgO(113) and Ag(113) diffraction peaks of the MgO:Ag sample. Units are in $2k_0 = 2/\lambda_0$.

The cluster size was estimated from the width of the (002) and (113) diffraction peaks using the Scherrer formula [84]

$$p = K_w \cdot \frac{\lambda_0}{\Delta\vartheta \cos \vartheta} = K_w \cdot \frac{2}{\Delta k\sqrt{h^2 + k^2 + l^2}} \qquad (35)$$

Here p is the precipitate size, defined as the cube root of the precipitate volume. Furthermore, λ_0 is the wavelength (1.540560 Å), θ is the position of the (hkl) diffraction peak, $\Delta\theta$ is the FWHM of the same diffraction peak, and Δk is the FWHM of the peak in reciprocal space in m^{-1}. The half-width Scherrer constant K_w depends on the shape of the nanocluster and—if the clusters are not spherically shaped—also on the *(hkl)* index of the diffraction peak [84]. The right-hand side of Eq. (35) is derived from Bragg's law, $\lambda = 2d_{hkl} \sin(\vartheta)$ and $k \equiv 1/d_{hkl}$. Here $d_{hkl} = d/\sqrt{h^2 + k^2 + l^2}$ where d is the lattice parameter of the material. k is the diffraction vector in reciprocal space and is commonly expressed in units of $2k_0 \equiv 2/\lambda_0$. The results of the Scherrer analysis are shown in Table 1. From the X-TEM results to be discussed, it was found that the shape of the Au clusters is spherical and the shape of the Ag clusters is octahedral. The corresponding values for the Scherrer constant K_w are listed in Table 1 and obtained from [84]. The cluster sizes were derived from the (113) diffraction peaks rather than from the (002) diffraction peaks, because the diffraction peaks of Au and Ag are better resolved from the MgO diffraction peak in the (113) case when compared to the (002) case. Therefore, the cluster sizes as obtained from the (113) are more accurate. From the positions of the Au(002) and Ag(002) peaks, the mean lattice parameter of the nanoclusters perpendicular to the surface was also calculated. It was found that the Au and Ag lattices are extended by 0.4% and 0.9%, respectively, in the direction perpendicular to the surface. This is probably caused by coherency strains, since the lattice parameters of Au and Ag are smaller than the lattice parameter of MgO (see Table 1). Calculation of the tensile strain from the (113) diffraction peaks yields the same result. The MgO:Au and MgO:Ag samples were analyzed with X-TEM. In the MgO:Au specimen, a band of Au clusters was found in a depth range of about 140–300 nm below the surface. Figure 14 shows a typical size distribution of the spherically shaped Au clusters. The size of the gold clusters varies from 2 to 14 nm with a mean size of 4.6 nm (average obtained from 40 clusters). Figure 15 shows a high-resolution TEM image of a few gold clusters. The cube-on-cube orientation relationship is clearly observable as the lines of the MgO lattice spacing continue in the Au nanocluster. X-TEM analysis of the MgO:Ag sample yields a depth range of 180–350 nm below the surface, a size distribution of 6–24 nm, and a mean cluster size of 11 nm (average obtained from 75 clusters). Here it was observed that the clusters have octahedral shapes except for the smallest clusters, which have spherical shapes.

In determining the mean cluster size, the Scherrer XRD method and the TEM observations are in excellent agreement. In the case of the Ag clusters, the Doyle optical method underestimates the cluster size, which might be attributed to additional broadening due to interface effects. Kreibig et al. [86] have shown that embedding of Ag nanoclusters in various matrices leads to additional broadening

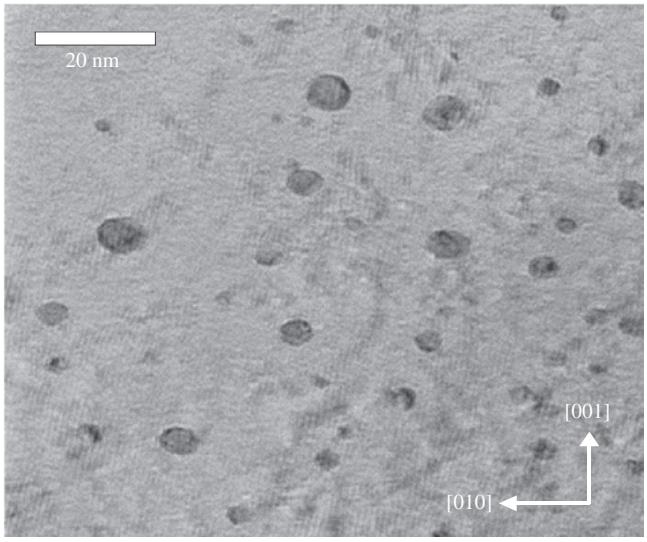


Figure 14. Bright-field TEM image of a part of the Au implanted layer showing a typical size distribution of spherically shaped Au clusters.

of the absorption peaks by a factor of 2 to 4 when compared to absorption peaks of Ag nanoclusters that are not embedded in a matrix, where the additional broadening is dependent on the embedding medium. Contemplating Eq. (34), additional broadening of $\Delta\lambda_{1/2}$ leads to an underestimation of the cluster size. XRD and TEM do not suffer from this artifact. However, both the Doyle method and the Scherrer method may suffer from the effect of size distribution, which, in general, also induces additional broadening.

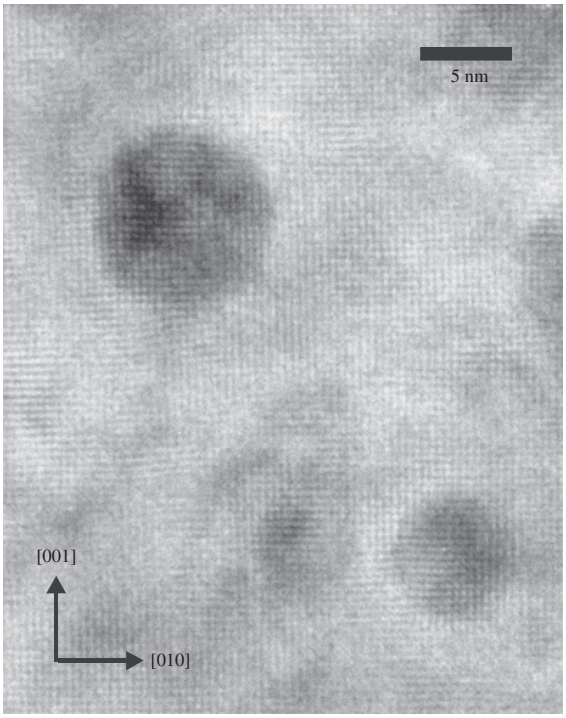


Figure 15. HRTEM image of a few gold clusters in a cube-on-cube orientation relationship with the embedding MgO matrix.

Nevertheless, the results shown in Table 1 show that the Scherrer method still accurately predicts the mean cluster size as observed by means of TEM. It was found here that of these three methods to determine the cluster size, the Doyle optical method is the least accurate. This is caused by the asymmetry of the absorption peaks, inaccurate background subtraction, and systematic underestimation due to interface effects. Scherrer XRD and X-TEM are both accurate methods to determine the mean cluster size, where X-TEM can also provide a good indication of the cluster size distribution. Considering that optical measurements take typically 10 minutes, XRD measurements a few hours, and X-TEM measurements about 3 days (including specimen preparation), Scherrer XRD is probably the best compromise. However, if one is interested mainly in the optical properties important for applications, optical absorption studies are indispensable. As shown in Figure 8 the absorption peak is a factor of 5 higher than in the case of RT implantations and the position of the peak is shifted to shorter wavelengths, $\lambda = 550$ nm. Moreover, in the case of the 1270 K implantation an extra peak at $\lambda = 800$ nm was observed. Assuming that the latter absorption peak results from the nanoclusters formed in the cavity layer the shift in the resonance wavelength can be explained by the elongated shape of the nanoclusters [87]. The photoabsorption measurements are summarized in Tables 2 and 3.

Besides Ag and Au, also copper ion implantation in MgO was studied. Two epipolished (100) MgO samples of size $1 \times 10 \times 10$ mm were implanted with copper ions at energy of 50 keV, sample 1 with a low dose of 1.0×10^{15} ions cm^{-2}, sample 2 with a high dose of 1.0×10^{16} ions cm^{-2}. In sample 2, nanovoids were present prior to copper implantation. These nanovoids were created by 30 keV implantation of 1.0×10^{16} He ions cm^{-2} and subsequent annealing at 1350 K [78]. After copper ion implantation, the samples were annealed isothermally and stepwise in air for 30 minutes, at temperatures of 550, 750, 1000, 1250, and 1350 K. After each annealing step, the samples were analyzed with the PBA technique using a 0–25 keV variable energy positron beam. After the 1250 K and 1350 K annealing step, the 1.0×10^{16} Cu cm^{-2} implanted sample 2 was analyzed using RBS-C. Finally, sample 2 was also analyzed with X-TEM after the 1350 K annealing step. A reference sample was present that underwent the same treatment as sample 2, except that no pre-existent cavities were present prior to copper ion implantation. The X-TEM result of sample 2 was

Table 2. Resonance wavelengths λ_r and half-maximum wavelengths (FWHM) $\Delta\lambda_{1/2}$ obtained from the photoabsorption measurements carried out in this study.

Implantation	Annealing T	Mean λ_r	$\Delta\lambda_{1/2}$
Dose (cm^{-2}), temperature (K)	(K/hour)	(nm)	(nm)
10^{16}, RT	1270/0.5	572	165
10^{16}, RT	1270/15	610	181
10^{16}, RT	1370/3	612	154
3×10^{16}, 1000	—	525	129
10^{16}, 1170	—	539	106
3×10^{16}, 1270	—	547	84
		818	101

Table 3. Mie resonance wavelengths λ_r and half maximum wavelengths (FWHM) $\Delta\lambda_{1/2}$ obtained from the optical absorption measurements.

Implantation	Annealing	Mean λ_r	$\Delta\lambda_{1/2}$	r_{cav}
dose (cm^{-2}), T (K)	T (K), time (hour)	(nm)	(nm)	(nm)
10^{16}, RT	1270, 0.5	572	165	1.46
10^{16}, RT	1270, 15	610	181	1.51
10^{16}, RT	1370, 3	612	154	1.79
3×10^{16}, 1000	—	525	129	1.57
10^{16}, 1170	—	539	106	2.02
3×10^{16}, 1270	—	547	84	2.62
		818	101	4.88

Note: $r = \frac{v_f}{\Delta\omega_{1/2}} = \frac{v_f}{2\pi c}\frac{\lambda^2}{\Delta\lambda}$

compared to an X-TEM analysis of sample 3. Sample 3 was implanted with 1.0×10^{16} Cu ions cm^{-2} at energy of 50 keV and was subsequently annealed only at 1250 K for 30 minutes. Sample 3 was used in another experimental series and the annealing took place in a reducing 5% H_2/95% Ar environment. However, the presence of a reducing environment only affects the relative areas of the (100) and (111) facets of the copper precipitate [88]. The sample treatment is listed in Table 4.

The PBA results for samples 1 and 2 are shown in Figure 16. The *S*-parameter corresponds to annihilation of positrons with conduction or valence electrons and is a strong indicator of open volume defects. On top of the figure a calculated depth distribution is shown of the copper ions after implantation and the depth distribution of implantation associated defects. These distributions were calculated with the TRIM code [81]. For comparison between the PBA data and the calculated TRIM data, the *S*-parameter has been plotted as a function of the depth in the material rather than as a function of positron energy. The depth has been calculated using the empirical relationship $\bar{z} = (\alpha/\rho)E^n$ with $\bar{z}$ the mean implantation depth in cm, E the positron energy in keV, ρ the density in g cm^{-3}, and α and n the empirical constants equal to 3.6×10^{-6} and 1.62, respectively [51].

It can be seen clearly that sample 1 (dose 1.0×10^{15} Cu cm^{-2}) results in a single layer at a depth of 0–40 nm under the surface with a quite high *S*-parameter value. These high values can only be obtained with large vacancy-type defects (size > 2–5 nm), where positronium can be formed. In this layer, nanovoids and presumably copper nanoprecipitates are present. In Figure 16, the same data are shown for sample 2 (dose 1.0×10^{16} Cu cm^{-2}). Here it can be observed that there are two peaks in the *S*-parameter graph, indicating that there are two layers of vacancy-type defects in the material, separated by a less defective layer. The formation of these two layers with vacancy-type defects does not depend on the presence of pre-existent cavities. The same PBA results were obtained for a reference sample that underwent the same treatment as sample 2 (implantation of 10^{16} copper ions cm^{-2} and annealing isothermally in air at temperatures of 550–1350 K), except for the fact that no preexistent cavities were present prior to copper ion implantation.

In both samples 1 and 2, the defects are located deeper in the material than predicted by TRIM. This might be the result of the long-range Coulomb interactions that are typical for ionic solids such as MgO. Furthermore, it can be observed that the pre-existent cavities present in sample 2 have been "washed out" by the copper implantation. The defects created by the copper implantation itself are thus dominant to the pre-existent cavities. The intermediate layer can be explained by assuming the presence of copper atoms, which decreases the *S*-parameter signal. An analysis of these PBA results has been performed by employing the positron beam analysis fitting program VEPFIT [49]. For sample 1, a three-layer model is used: first a layer with 0–40 nm depth representing the implantation area, a second layer with 40–130 nm depth representing a "tail" of defects stretching into the material, and the bulk as the third layer. For sample 2,

Table 4. Overview of sample treatment and main observations for three (100) MgO samples.

	Sample 1	Sample 2	Sample 3
Pre-existent cavities	—	1.0×10^{16} ^{3}He ions cm^{-2} and 1350 K annealing [76]	—
Implantation dose	1.0×10^{15} Cu ions cm^{-2} at an energy of 50 keV	1.0×10^{16} Cu ions cm^{-2} at an energy of 50 keV	1.0×10^{16} Cu ions cm^{-2} at an energy of 50 keV
Isothermal annealing	for 30 minutes at 550, 750, 1000, 1250 and 1350 K in air	for 30 minutes at 550, 750, 1000, 1250 and 1350 K in air	for 30 minutes at 1250 K in H_2/Ar gas
Positron beam analysis	• one single layer with vacancy-type defects	• implantation created defects are dominant to pre-existent cavities • two separated layers with large vacancy-type defects	—
RBS-C	—	• diffusion of copper into the bulk during annealing • copper atoms reach the cavities layers at depths over 100 nm • copper atoms epitaxial to the MgO lattice	—
X-TEM		• rectangular voids • no presence of copper precipitates	• copper precipitates next to small rectangular voids

Note: Not shown is a reference sample that underwent the same treatment as sample 2, except for the fact that no pre-existent cavities were present prior to copper ion implantation.

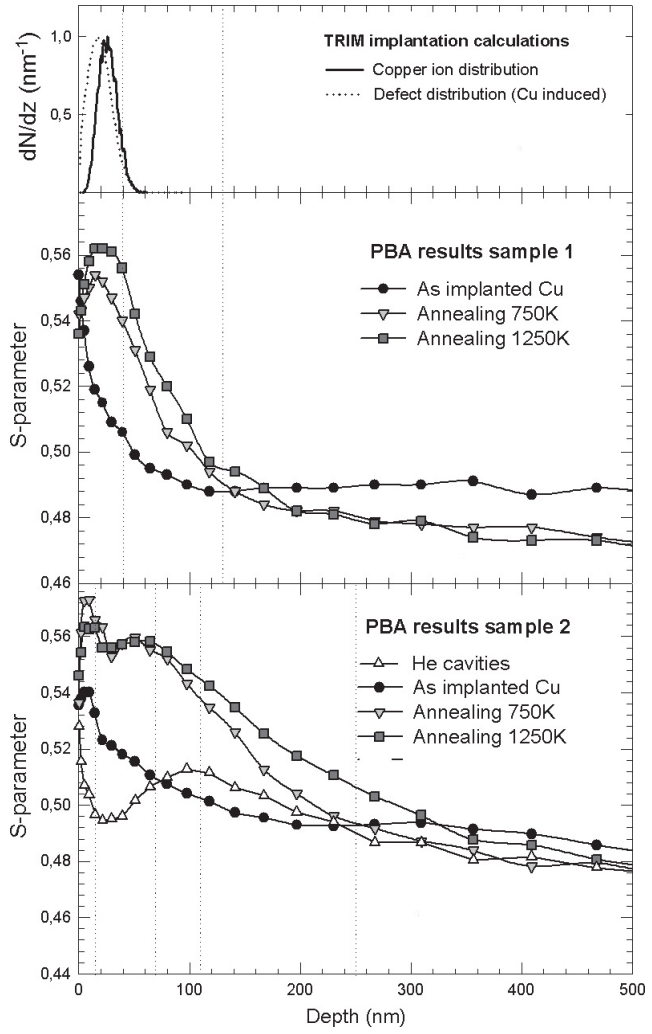


Figure 16. TRIM calculations and positron beam analysis results. Top: TRIM predicted depth resolved distribution of copper atoms and created defects after implantation. Center: *S*-parameter as a function of depth for sample 1 (dose 1.0×10^{15} Cu cm^{-2}) as implanted and after annealing at 750 and 1250 K. Bottom: *S*-parameter as a function of depth for sample 2 (dose 1.0×10^{16} Cu cm^{-2}) before implantation (with pre-existent cavities), as implanted with copper ions, and after annealing at 750 and 1250 K.

a five-layer model is used. The two peaks in the spectrum can only be ascribed to two highly damaged areas in the material, separated by a less defective layer. These are the first three layers, with boundaries at 0–15, 15–70, and 70–110 nm. Here again a "tail" of defects in the material is present from 110 to 250 nm and the fifth layer is the bulk. The *S*-parameters for each layer have been fitted and these values are plotted versus the sample treatment in Figure 17. In the five-layer model used for sample 2, the highly damaged layers 1 and 3 behave quite similar during the annealing procedure, while the intermediate layer 2 remains stable. This implies that the vacancy-type defects in layer 1 and 3 aggregate during the annealing process, and that these big vacancy-type defects are not present in the intermediate layer.

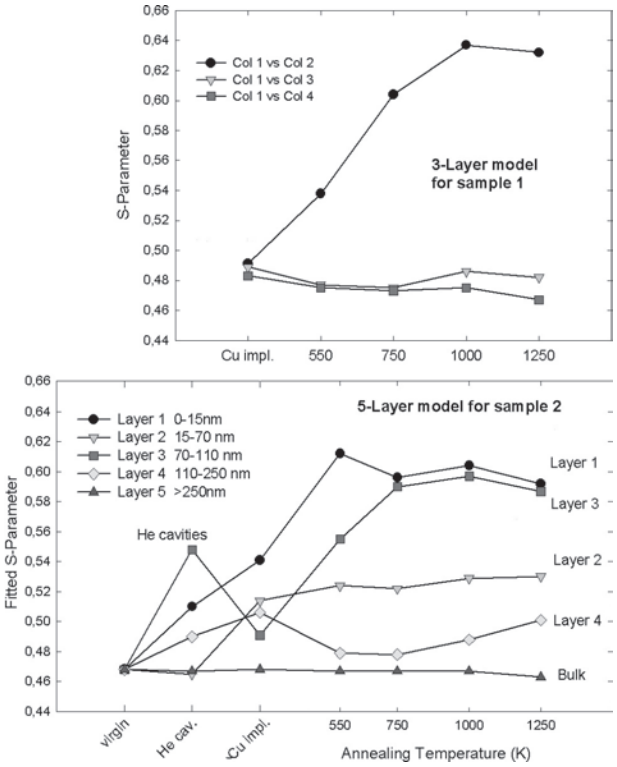


Figure 17. VEPFIT modeled *S*-parameters in each layer versus sample treatment. The "virgin" point is for clean unused MgO samples. Top: three-layer model for sample 1 (see text). Bottom: five-layer model for sample 2 (see text).

The high dose copper implanted sample was analyzed with RBS-C to observe the depth resolved distribution of the implanted copper atoms; the result is shown in Figure 18 and listed in Table 4. The first observation is that the copper atoms diffuse into the bulk during the annealing procedure. The total area under the signal is approximately equal

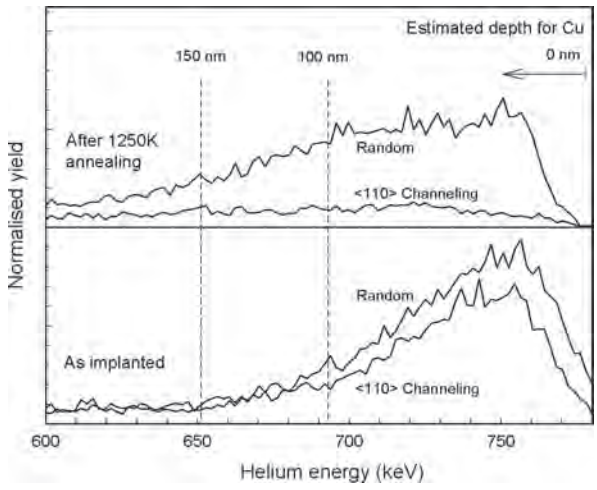


Figure 18. RBS-C results for the 1.0×10^{16} Cu ions cm^{-2} implanted sample 2, "as-implanted" and after the 1250 K annealing step. For both stages in the sample treatment, a random alignment (RBS) and a ⟨110⟩ alignment (channeling) are shown.

for the "as-implanted" and the "1250 K annealing" cases, indicating that the copper does not segregate at the surface. Combining this with the PBA results, the copper atoms have certainly reached the "deep" layer of nanocavities at depths over 100 nm. The second observation is that after the 1250 K annealing step, the copper atom lattice is epitaxial with respect to the MgO host lattice, as the signal is reduced to 20% when a channeling configuration is used. The X-TEM analyses that will be discussed show that copper nanoprecipitates are indeed epitaxial to the MgO host lattice.

From the previous discussion, it becomes clear that copper implantation and subsequent annealing cause a high concentration of defects that aggregate into bigger nanovoids. At the same time, the copper atoms diffuse into the bulk during the annealing process, with the copper atom lattice epitaxial to the MgO lattice. The question now arises as to what configuration the copper atoms are present during the various stages of the sample treatment. To answer this question, sample 2 was analyzed using X-TEM after the 1350 K annealing step. This result was compared to the X-TEM analysis of sample 3 that was implanted with the same dose of 1.0×10^{16} copper ions cm^{-2} but was annealed only once after implantation: at 1250 K for 30 minutes in a reducing environment. The results are displayed in Figures 19 and 20 and listed in Table 4.

Considering the X-TEM results of sample 2 (Fig. 20), it is clear that large cavities (size up to 20 nm) are present at a depth of 200–400 nm and that they are rectangularly shaped. The PBA results do not indicate large cavities at depths over 250 nm, possibly because of the low density of these defects. The rectangular shape of the cavities is probably due to the fact that the surface energy is lowest for the (100) face of MgO: the surface energy of the (100) face is calculated to be 1.2 J m^{-2}, and the surface energy of any other (faceted) MgO surface to be at least 1.8 J m^{-2} [82]. In sample 3 which was annealed at 1250 K in a reducing environment, nanovoids and copper nanoprecipitates coexist and do not form common clusters. Presumably there is a bias for copper

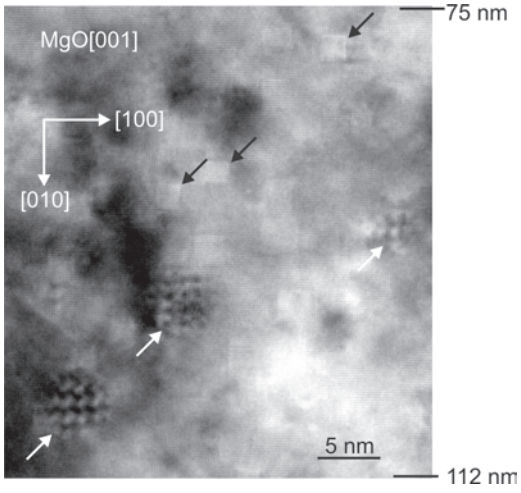


Figure 19. X-TEM result of sample 3 after the 1250 K annealing step. Copper precipitates (white arrows) are present next to small rectangular voids (black arrows) of size 2–3 nm.

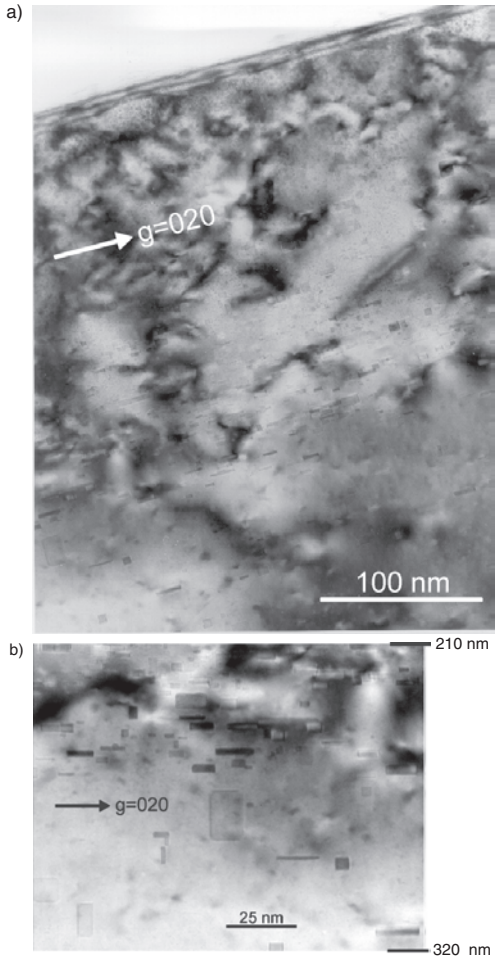


Figure 20. X-TEM result of a high dose implanted sample after one 1350 K annealing step. (a) Overview with surface (depth up to 400 nm). (b) Detail of rectangular cavities.

interstitial-type defects and vacancy-type defects to recombine. Furthermore, the copper precipitates seem to have a spherical shape but in fact the Cu//MgO interface consists of (100) and (111) facets [88]. In the 1350 K annealed sample 2 depicted in Figure 20 copper nanoprecipitates can no longer be found. Possibly the copper atoms are trapped at the internal surface of the nanovoids, although this cannot be put with certainty. An order-of-magnitude calculation on the X-TEM and RBS results yields that in the high-density void areas with a void fraction of 10^{-3}, less than 10% of the internal surfaces can be covered with copper atoms. The copper atoms might also be completely dissolved in the MgO lattice, replacing certain Mg atoms. This is in some concordance with results presented by Zimmerman et al. [82] on copper precipitates in MgO. These workers observed that the absorption peak disappeared after annealing at 500 K for 1 h, assuming dissociation of the copper precipitates. In this work, the dissociation of the copper precipitates occurs only after high temperature annealing. A feasible explanation for this discrepancy is that the formation and dissociation of metal precipitates depends on the annealing time and temperature as well as on the presence of a (non)reducing environment. The optical absorption spectroscopy technique

used by Zimmerman et al. was not used in this work because the implantation doses used are too low to produce significant absorption peaks.

The major conclusions that can be drawn from the above discussion are:

- A copper implantation dose of 1.0×10^{15} ions cm^{-2} results in a single layer of nanovoids; a copper implantation dose of 1.0×10^{16} ions cm^{-2} results in two layers of nanovoids. The defects induced by 50 keV implantation of 1.0×10^{16} cm^{-2} copper ions dominates the presence of pre-existent nanovoids at depths over 100 nm.
- The copper atoms diffuse into the bulk during annealing, reaching the nanovoid area at depths over 100 nm.
- Copper nanoprecipitates are epitaxial to the MgO host lattice and can coexist with nanovoids.
- Nanovoids formed in MgO are rectangularly shaped.
- After annealing at high temperature and/or for long periods of time the copper precipitates dissociate and resolve into the MgO host lattice

4. POSITRON CONFINEMENT IN EMBEDDED LITHIUM NANOCLUSTERS

4.1. Introduction

Metal and semiconductor nanoclusters are of interest because of their linear and nonlinear optical properties but also because of their electronic features as quantum dots [89–91]. The optical and electronic properties are strongly dependent on the nanocluster size, which is due to quantum size effects such as the confinement of surface plasmons and the discretization of energy levels below a certain cluster size. The engineering of quantum dots offers great opportunities for the future as the material properties, such as the bandgap and melting temperature of semiconductor clusters, can be tuned by varying the cluster size [92].

In this field positron confinement is a newly investigated type of quantum confinement in nanoclusters and can serve as a powerful tool to investigate the electronic structure of nanoclusters. Positrons are known as sensitive self-seeking probes of defects in materials and can provide detailed information on the electronic structure of defects or bulk materials [93, 94] such as the morphology of Fermi surfaces [95, 96]. The concept of positron confinement was introduced by Nagai et al. [97]. Under certain conditions, nanoclusters act as a potential well to positrons (i.e., the wave function of the positron is spatially confined within the nanocluster in three dimensions). A positron confined in a nanocluster will annihilate there and will thus provide unique information on the electronic and crystal structure of the nanocluster, which can be quite different from the bulk material of the atoms constituting the nanocluster. Previously, positron confinement in Cu nanoclusters embedded in Fe was investigated [97]. By analysis of the S and W parameters of the Doppler broadening curves (see Section 2) these authors found that, after thermal aging, approximately 90% of all positrons annihilate in the Cu nanoclusters while Cu makes up only 1.0 wt% of the Fe:Cu composite material. The positron confinement was confirmed by coincidence Doppler broadening (CDB) measurements and was ascribed to a difference in positron affinity between Cu and Fe of 1.0 eV. If the positron affinity of the cluster material is lower than the positron affinity of the host material, the cluster will act as a potential well to positrons, as will be explained in Section 4.2.

In this section positron confinement is discussed in Li nanoclusters embedded in monocrystalline MgO. In order to study the material properties of nanoclusters, it can be useful to embed nanoclusters in stable and inert ceramics such as MgO, which can be achieved by ion implantation and subsequent annealing [98]. It is expected that the wide bandgap material MgO (7.8 eV) does not interact with the electronic structure of metal and semiconductor clusters [98]. It is also optically transparent in a large frequency band, which facilitates optical studies. Furthermore, the high melting point of MgO (3070 K) allows a study of phase transitions of nanoclusters. The optical and structural properties of Li nanoclusters embedded in MgO have been investigated by photoabsorption spectroscopy, transmission electron microscopy, and electron diffraction [99]. It was found that small Li clusters (<20 nm) adopt the unusual fcc crystal structure while large Li clusters (>40 nm) adopt the regular body-centered cubic (bcc) crystal structure. This is a so-called extrinsic size effect: small Li clusters adopt a different phase in order to fit better in the MgO matrix. The lattice parameter of fcc lithium is closer to that of MgO, which reduces the formation energy of the MgO//Li interface. The fcc phase for bulk Li is otherwise only observed at low temperatures and high pressures [100, 101]. A similar effect was reported in the work mentioned previously [97] where the crystal structure of Cu nanoclusters embedded in Fe is expected, on the basis of lattice parameter similarity, to be bcc rather than fcc.

The nanoclusters shown in this section were created by ion implantation and subsequent annealing. It was shown previously in [102] that Doppler broadening PBA and 2D-ACAR measurements show a very high fraction of positron annihilations in Li nanoclusters, indicating possible positron quantum confinement. In the following we will first discuss the concepts of positron affinity and the positronic potential well. Furthermore, the positron affinities of MgO and fcc and bcc lithium are calculated in order to determine the drop of the positronic potential at the MgO//Li interface. The confinement of positrons in Li nanoclusters is discussed using the experimental and computational results.

4.2. Theoretical Methods and Concepts

As explained in Section 2 annihilation of positrons with electrons in solids yields information on the momentum distribution of these electrons. The electronic momentum distribution is reflected in the Doppler broadening of the 511 keV annihilation peak. Positron annihilation with low momentum valence or conduction electrons results in a small Doppler shift, contributing to the center of the peak. Annihilation with high momentum core electrons results in a large Doppler shift, contributing to the wings of the 511 keV annihilation peak. The shape of the 511 keV peak is characterized by the so-called S and W parameters [8]; see also section 2. When combined with a slow positron beam, this

method is called Doppler broadening positron beam analysis. The S and W parameters can be considered as a specific bulk property for every material. In the case of composite materials (e.g., metallic nanoclusters embedded in insulating oxides), the overall annihilation distribution can be modeled as a linear combination of the annihilation distribution of each of the composing materials. The S parameter of the composite material then equals

$$S_{\text{comp}} = f_{\text{cluster}} S_{\text{cluster}} + f_{\text{host}} S_{\text{host}} \tag{36}$$

where S_{cluster} and S_{host} are the bulk S parameters of the nanoclusters and the embedding material, respectively. f_{cluster} and f_{host} are the fractions of annihilation in the clusters and in the embedding host material ($f_{\text{cluster}} + f_{\text{host}} = 1$). Here it is assumed that the composite material is defect-free and that positrons do not annihilate at the cluster–host interface. Of course, the principle of superposition of annihilation distribution can be applied to other positron annihilation methods as well, including CDB and 2D-ACAR. Equation (36) applies to a "bulk" composite material. In the case of, for example, an ion-implanted layer investigated using a positron beam, the S parameter obtained at certain positron implantation energy consists of contributions from the surface and various layers,

$$S(E) = f_{\text{surf}}(E)S_{\text{surf}} + f_1(E)S_1 + f_2(E)S_2 + ::: + f_n(E)S_n \tag{37}$$

Here the material consists of n layers and the distribution of the fractions ($f_{\text{surf}} + f + ::: + f_n = 1$) depends on the positron energy E. Please note that Eq. (36) refers to defect (or material) fractions while Eq. (37) refers to layer fractions. Formulae analogous to Eq. (36) are valid for the W parameter. Fitting methods such as VEPFIT and SWAN [103] exist, which yield accurate values for the S and W parameters of the various layers and the fractions of Eq. (37) positrons annihilating in these layers as a function of positron implantation energy. This is achieved by solving the time-averaged diffusion equation [8]. Once the S parameter of a certain layer has been found using the VEPFIT code, Eq. (36) can be used in order to analyze further the defect fractions within that layer.

Assuming that the nanoclusters are homogeneously distributed, the annihilation fractions f_{cluster} and f_{host} are determined by the diffusion length of the positrons, the concentration of clusters, and the difference in positron affinity between both materials. Although the overall concentration of atoms constituting the clusters in the host material is usually small, the annihilation fraction in the clusters can be high with respect to the annihilation fraction in the host material when certain conditions with respect to the positron affinities are met.

Since the Fermi levels of the materials in contact become equal [104], the positron affinity is defined as [105]

$$A_+ = \mu_- + \mu_+ = -(\varphi_- + \varphi_+) \tag{38}$$

where μ_- and μ_+ are the electron and positron chemical potential, and φ_- and φ_+ are the electron and positron work functions, respectively. Values for the positron affinity are commonly expressed in eV and are negative for most materials. At the interface between two materials, positrons will encounter a potential step that is equal to the difference in the positron affinities of both materials [105]

$$\Delta E_+^{AB} = A_+^A - A_+^B \tag{39}$$

A necessary condition for positron confinement in a cluster (A) embedded in a host (B) is that $\Delta E_+^{AB} > 0$. Then the cluster represents a potential well for positrons. However, there must be at least one bound state in such a well, which determines the minimum size of the cluster. Assuming a spherical potential well this condition reads [105]

$$R > \frac{\pi\hbar}{\sqrt{8m_0\Delta E_+^{AB}}} \tag{40}$$

where R is the radius of the cluster and m_0 is the positron mass. In Figure 21, several possibilities for the positron potential energy as a function of the distance from the center of the nanocluster are plotted schematically. In Figure 21a, $\Delta E_+^{AB} > 0$. Here the nanocluster acts as a positronic potential well with depth ΔE_+^{AB}. Positrons can be confined in this nanocluster provided that condition (40) is satisfied. In Figure 21b, $\Delta E_+^{AB} < 0$. The nanocluster acts as a three-dimensional potential barrier and cannot confine positrons. Figure 21c shows a situation where the positron affinity of the nanocluster material is lower than the positron affinity of the host material, but positron confinement is unlikely as the positron is trapped at the interface. Often there is some lattice mismatch due to the different crystal structures of the cluster and the embedding matrix. At the interface, this may result in structural and electronic defects, which can act as a trap for positrons. In the case of, for example, a Si//SiO2 interface, there is evidence of preferential trapping at this interface [106].

Usually, the positron affinity is determined experimentally by measuring the electron and positron work functions

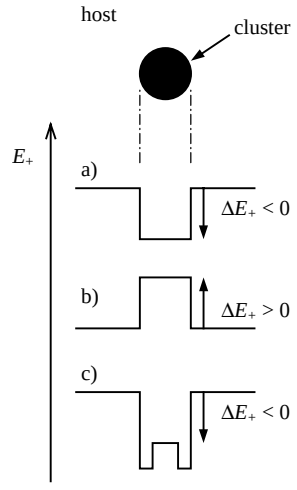


Figure 21. Diagram of the positron potential energy as a function of the distance from the center of the nanocluster. (a) Nanocluster acts as a positronic potential well with depth ΔE_+. (b) Nanocluster acts as a potential barrier with step size ΔE_+. (c) Positron trapping at defects at the cluster–host interface.

[107] [see Eq. (38)]. Alternatively, the positron affinity can be derived from the positronium formation potential ε_{Ps}, which can be determined experimentally as the negative of the maximum kinetic energy of positronium atoms ejected from the sample into vacuum. The positronium formation potential is related to the positron affinity as

$$\varepsilon_{Ps} = \phi_+ + \phi_- - 6.8\ \text{eV} = -A_+ - 6.8\ \text{eV} \qquad (41)$$

where 6.8 eV is the binding energy of the positron–electron pair constituting a positronium atom in vacuum. Two experimental time-of-flight studies on MgO [108] report positronium emission peak energies of 4.5 and 2.6 eV, respectively. Applying Eq. (41) then yields values for the positron affinity for MgO of −2.3 and −4.2 eV, respectively. Apart from the quantitative discrepancy, the interpretation of these results is complicated by the fact that the emission energy spectra are quite broad. Furthermore, the unknown binding energy of Ps at the MgO surface (MgO is too dense a material for Ps to be formed in the bulk) and the uncertainty about whether valence or conduction electrons participate in Ps formation obscure a direct translation to the positron affinity for MgO [108, 109]. Summarizing, determining positronium formation potentials at the surface of insulating oxides such as MgO is experimentally difficult, which results in the different values reported in the literature.

Therefore, we turn to *ab initio* calculations to determine the positron affinities. As the calculation of electron and positron work functions requires a careful treatment of the material's surface, we prefer to determine A_+ through the electron and positron chemical potentials [see Eq. (38)], the calculation of which is a standard procedure in "bulk" electronic structure computational methods. It should be noted that A_+ is a bulk characteristic of a material (see [94, 105] for details). For the calculation of the positron affinities and lifetimes in the materials studied [94] the linear-muffin-tin-orbital (LMTO) method was used within the atomic sphere approximation (ASA) [110, 111]. Within the framework of this method the crystal is divided into slightly overlapping spheres (called atomic spheres) centered at the positions of the nuclei. The sum of the volumes of the spheres is equal to the volume of the crystal. However, the rock salt structure of MgO is not as closely packed as the fcc and bcc structures considered for Li [112]. Therefore we had to introduce empty spheres [110] into the MgO lattice to describe properly the distribution of the electron and positron densities and potentials in the interstitial region. We considered the case where the empty spheres occupied 30% of the cell volume. In the following this case is referred to as MgO II, in contrast to MgO I where no empty spheres were included for comparison. Choosing the 30% filling with empty spheres leads to a reasonable overlap between neighboring Mg–O spheres in the MgO lattice [110]. The lattice constants a_0 used in our calculations are collected in Table 5. As for the positron calculations, in Li we used the parametrizations of the positron correlation potential and enhancement factor given by Boronski and Nieminen [113] within the framework of the local density approximation (LDA). The corresponding results will further be denoted as LDA, whereas the results obtained using the generalized gradient approximation [114] will be marked GGA. In the case of MgO we also employed the GGA, but instead of the LDA we made use of the so-called semiconductor (SM) and insulator (IM) models [115], which take into account incomplete positron screening in nonmetallic systems (i.e., systems having a bandgap; MgO is an insulator). A high frequency dielectric constant $\varepsilon_\infty = 3.0$ was employed in the SM and IM calculations. This constant was derived from the optical constants of MgO [116] in the visible spectrum $(n, k) = (1.73, 0.00)$ using the relationship $\varepsilon_\infty = n^2 - k^2$. Finally, the exchange-correlation potential for electrons parametrized according to Von Barth and Hedin [117] was used in the calculations of the electronic structure of MgO and Li.

Table 5. Structural data.

	Structure	a_0 (Å)	At. volume (Å^3)
MgO	rocksalt	4.212[a]	—
Li	fcc	4.40[a]	21.3
Li	bcc	3.51[b]	21.6

[a] From "Handbook of Chemistry and Physics," 56th ed., p. B-214. CRC Press, Boca Raton, FL, 1975.
[b] From [112].

For completeness, the positron lifetime (τ) in defect-free MgO and Li was calculated employing the well-known formula [94]

$$1/\tau = \pi r_0^2 c \int n_-(\mathbf{r}) n_+(\mathbf{r}) \gamma[n_-(\mathbf{r})] d\mathbf{r} \qquad (42)$$

Here, n_- and n_+ denote the electron and positron density, respectively, and γ stands for the enhancement factor describing the pile up of electrons around a positron (r_0 is the classical radius of the electron and c the speed of light). The form of γ within the LDA, GGA, SM, and IM approaches was used in the corresponding calculations. Other computational details are described in [118].

4.3. Experiments

Epi-polished monocrystals of MgO(100) of size $10 \times 10 \times 1$ mm^3 were implanted with 1.0×10^{16} cm^{-2} ^{6}Li ions at an energy of 30 keV. After the implantation the crystals were annealed in air in steps from room temperature to 1200 K. The annealing temperatures were 550, 750, 950, 1100, and 1200 K for periods of 30 min. Photoabsorption spectroscopy was used to detect Mie plasmon resonance due to the presence of lithium nanoclusters. Furthermore, the neutron depth profiling (NDP) technique, using the nuclear reaction $^6\text{Li}(n, \alpha)^3\text{H}$, was applied to determine the depth profile of ^{6}Li after the annealing steps. The results of these two methods have already been discussed in an earlier paper, which focused on the defect evolution during the annealing procedure [102]. In the present work we address the confinement of positrons in lithium nanoclusters, and therefore only the positron experimental techniques applied to the samples are presented.

After ion implantation and after each annealing step, the samples were analyzed with Doppler broadening PBA using a monoenergetic 0–30 keV positron beam. The energy resolution of the PBA setup is 1.2 keV. After the 950 K annealing step, one of the samples was also analyzed with 2D-ACAR.

The 2D-ACAR setup of the Anger-camera type is coupled to a high-intensity positron beam with a flux of 8×10^7 e^+ s^{-1}, allowing depth-selective 2D-ACAR [119]. The angular resolution of this system is 1.1×1.4 mrad^2. The sample measured with 2D-ACAR was identical to the other samples (which was ascertained by means of photoabsorption spectroscopy and PBA), apart from a 15-nm-thick layer of Al deposited on the sample surface after the 950 K annealing step to prevent buildup of electric charge. The 2D-ACAR distribution was collected at a positron implantation energy of 4 keV, which corresponds to a mean positron implantation depth at the center of the layer containing the lithium nanoclusters. This energy was chosen on the basis of depth profiling performed on the same sample by means of Doppler broadening PBA. Furthermore, an MgO bulk 2D-ACAR distribution was collected for reference purposes from an as-received sample of MgO(100). The use of a conventional ^{22}Na source in this latter measurement renders any surface effects negligible.

4.4. Results

4.4.1. Positron Beam Analysis

Figure 22 shows the *S* and *W* parameters as a function of positron implantation energy after lithium ion implantation and after annealing at the indicated temperatures. The solid lines represent the result of simulations performed with the VEPFIT code, as will be discussed. The *S* parameter shows a spectacular increase in the lithium ion implantation range (corresponding to a positron energy of 3–4 keV) after annealing at temperatures of 750 and 950 K. At higher annealing temperatures the *S* parameter in the ion implantation layer drops considerably and after annealing at 1200 K the peak almost disappears in this range. At larger depths (corresponding to positron implantation energies of 7–15 keV in Fig. 22), the *S* parameter drops below the bulk value of MgO (i.e., 0.468) after annealing at 1200 K. Such a drop of the *S* parameter below the bulk value of MgO as a result of ion implantation and subsequent annealing has not been observed before and we attribute this effect to positron annihilation in small Li-related defects, as will be discussed in more detail.

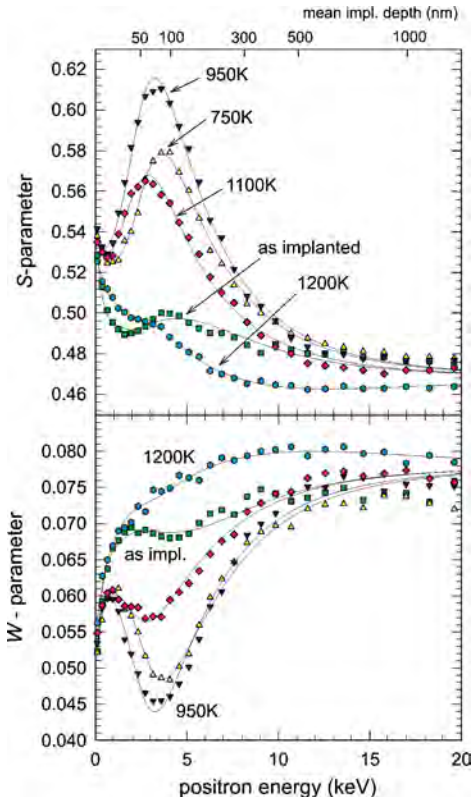


Figure 22. *S* and *W* parameters as a function of positron energy. The solid lines are the result of VEPFIT modeling.

The experimental *S* and *W* parameter curves shown in Figure 22 have been fitted using the VEPFIT code in order to gain more insight into defect evolution and in order to derive the contributions of the various layers to the annihilation distribution. A four-layer model was used. The principal defect types and the corresponding diffusion lengths that were used as input to the VEPFIT code are listed in Table 6. The choice for this model is mainly based on defect analysis performed on previously ion implanted MgO samples that were analyzed with techniques such as transmission electron microscopy, positron annihilation spectroscopy, and neutron depth profiling [120, 121]. In particular, there is a subsurface layer containing the ion-implanted species and the main implantation defects (layer 2). Dislocation loops are formed at smaller depths in the MgO top layer (layer 1) and below the ion implantation layer [120]. The presence of dislocation loops in MgO does not affect the *S* parameter significantly, but it does shorten the diffusion length compared to the bulk values. In the layer below the ion implantation layer (layer 3) there are not only dislocation loops but also a tail of implanted species as observed by NDP [102], possibly because of channeling effects. It was estimated that about 6% of the implanted species end up in this range. Due to the low atomic concentration of Li in this layer, we expect small Li-related defects rather than metallic Li clusters (see the discussion on the defect evolution of layer 3 in Fig. 22). The fourth layer is the MgO bulk. In order to monitor the defect evolution, the *S* and *W* parameters of layer 2 (ion implantation layer) and layer 3 (ion implantation tail) were fitted. The layer boundaries determining the position of the ion implantation layer have been allowed to relax as the *S* parameter peak shifts to somewhat lower positron energies when the annealing temperature increases (compare, for example, the 750 and 1100 K *S* parameter curves in Fig. 22). This might be due to a recombination of implanted lithium atoms and lithium clusters with vacancies

Table 6. VEPFIT model used to fit the *S* and *W* parameter curves shown in Figure 21.

	Depth (nm)	Diff. length (nm)	*S* par.	*W* par.	Description
Layer 1[a]	0–d_1	50	0.468	0.078	MgO top layer
Layer 2	d_1–d_2	15	fitted	fitted	MgO:Li
Layer 3	d_2–300	50	fitted	fitted	ion impl. tail
Layer 4	>300	100	0.468	0.078	MgO bulk

[a] In the case of the 1200 K annealing step, layer 1 was omitted ($d_1 = 0$).

and small vacancy clusters created by ion implantation. The main ion implantation damage is always located at slightly smaller depths than the range of the implanted ions.

The fitted curves are depicted as solid lines in Figure 22. The agreement with the experimental results is reasonable, considering the fact that the same model has been applied to all data sets. In the case of the S and W graphs corresponding to the 1200 K annealing step the MgO top layer was omitted since it was not possible to distinguish between the ion implanted layer and the MgO top layer. The fitted results yield boundaries d_1 and d_2 of the ion implanted layer of approximately 45 and 115 nm respectively, shifting to slightly lower values with increasing annealing temperature. The fitted values of S and W are presented in Figure 23b.

Figure 23 shows the S parameter plotted vs the W parameter. Every bulk material and every defect type has a characteristic S and W parameter. A few of these typical S, W combinations are indicated in Figure 23 with open circles and are called cluster points. The cluster points corresponding to pure MgO and bcc Li were determined experimentally from bulk samples of these materials. By plotting experimental or fitted S–W values in an S–W plot with, for example, the positron energy as the running parameter, the defect status at a certain depth in the material can be analyzed by means of comparison with the characteristic S–W cluster points for defects and bulk materials. In Figure 23a, the experimentally obtained S–W parameters are plotted with the positron energy as the running parameter. Figure 23b shows the S–W plot using the fitted S and W values for every layer with the annealing treatment as the running parameter. The main difference between the two plots is that in Figure 23a, all layers contribute to the experimental S–W points, while in Figure 23b the S–W parameters are layer-resolved with the aid of the VEPFIT code; that is, the S and W parameters corresponding to only one layer are shown [see also Eq. (37)].

Considering the defect evolution of layer 3 (ion implantation tail) in Figure 23b, it is clear that the S, W combination moves away from the bulk Li cluster point during the annealing procedure, while the NDP measurements show that a small fraction (approximately 6%) of the implanted Li atoms is present in this layer. Therefore, the Li atoms should be in an electronic state that is different from metallic bulk Li.

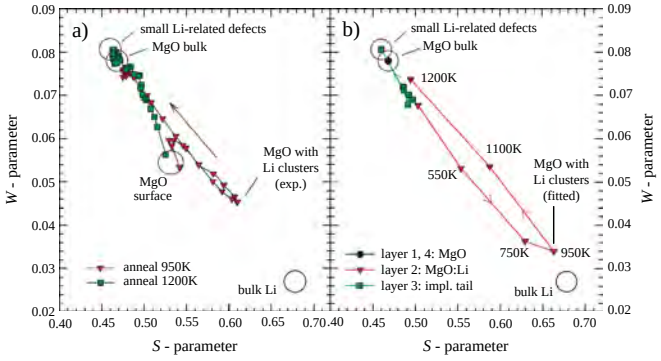


Figure 23. (a) SW plot with the positron energy as running parameter. The experimental data are plotted for the 950 and 1200 K annealing step only. (b) SW plot with the annealing treatment as running parameter. The fitted values are plotted for all four layers (see Tab. 6).

Research performed previously on Li-doped MgO crystals has revealed defect types such as Li_2O precipitates, isolated Li^+ ions, and so-called $[Li]^0$ centers [122]. These defects will yield S–W parameters that are quite different from bulk Li or bulk MgO. Another observation that can be made with respect to Figure 23b is that the S–W characteristic points corresponding to the ion implantation layer (layer 2) move roughly along the line connecting the MgO bulk cluster point and the bcc bulk Li cluster point. It is also clear that the experimental S–W point corresponding to MgO containing Li clusters after the 950 K annealing step in Figure 22a does not reach the fitted S–W combination of layer 2 (MgO containing lithium clusters) after the 950 K annealing step in Figure 23b. Due to the broadness of the positron implantation profile and positron diffusion processes, only a certain fraction of the positrons implanted at 3.5 keV (with a mean implantation depth that corresponds to the center of the implantation layer) annihilate in the MgO:Li layer. This is also clear from Figure 24 where the fraction of annihilations per layer is plotted as a function of positron implantation energy. It is found that 69% of the positrons with 3.5 keV implantation energy annihilate in MgO:Li (layer 2).

Calculation of the efficiency of positron confinement in Li nanoclusters using Eq. (36) requires numerical values for the S parameter of the composite MgO:Li material (i.e., the S parameter of layer 2), the S parameter of bulk lithium, and the S parameter of bulk MgO. The S parameter of the MgO:Li layer ($S_{MgO:Li}$) after the 950 K annealing step was found to be 0.662 by means of VEPFIT modeling (Tab. 6, Fig. 23b). The S parameters of bulk lithium and bulk MgO were measured using the same Doppler broadening PBA setup and were evaluated to be 0.678 for polycrystalline bulk bcc Li and 0.468 for monocrystalline bulk MgO. The measurement of the S parameter of bulk Li was carried out at a positron implantation energy of 27 keV. This energy is sufficiently high to make the contribution of the lithium oxide layer at the surface to the measured distribution negligible, which was ascertained by means of positron depth profiling. The S parameter for bulk lithium is much higher than the S parameter for bulk MgO. This is due to the rather

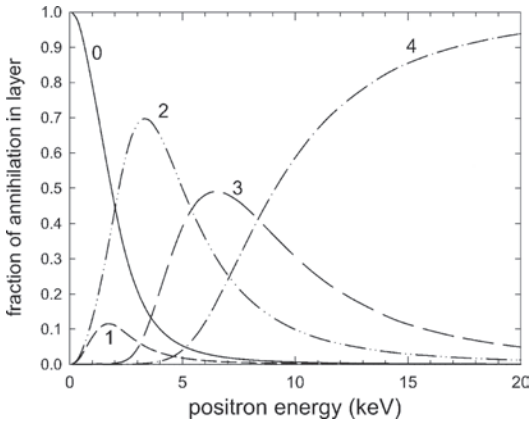


Figure 24. Layer fractions of the PBA sample as a function of the positron implantation energy. Annihilation fractions per layer were found by VEPFIT modeling of the PBA data obtained after the 950 K annealing step. The numbers indicate the layer number as listed in Table 6. Layer 0 is the surface (thermal and epithermal positrons).

small Fermi cutoff for metallic lithium of 0.58 a.u. [123] (equivalent to 4.2 mrad) while the window for the measurements of the S parameter is set at energies corresponding to ± 3.2 mrad. Consequently, most of the Doppler broadening curve falls within the window and the S parameter for lithium is therefore relatively high. This is also clear from Figure 25 where the experimentally obtained 511 keV annihilation peaks of bulk MgO, bulk bcc Li, and MgO:Li are displayed. The fraction of annihilations in lithium can be calculated rewriting Eq. (36) and using $f_{\mathrm{MgO}} + f_{\mathrm{Li}} = 1$

$$f_{\mathrm{Li}} = \frac{S_{\mathrm{MgO:Li}} - S_{\mathrm{MgO}}}{S_{\mathrm{Li}} - S_{\mathrm{MgO}}} \tag{43}$$

Here it should be realized that contributions from the surface and any other layers have already been filtered out by the VEPFIT code. Substituting the S parameters mentioned previously then yields a fraction of annihilations in lithium, f_{Li}, of $92 \pm 5\%$. This shows a very effective confinement of positrons in Li nanoclusters, which occupy only 1.3 at% of the MgO:Li layer. In the previous calculation, it is assumed that the S parameters for bcc Li and fcc Li do not differ significantly and that the S parameter of the lithium nanoclusters (without surrounding MgO) is similar to the S parameter of lithium bulk. A question that now arises is whether the positron is really confined within the Li nanocluster or whether it is trapped at the MgO//Li interface prior to annihilation.

4.4.2. 2D-ACAR

The 2D-ACAR (see Section 2) results are shown in Figure 26. Part (a) shows the anisotropic contribution [124] of the MgO bulk distribution (total number of counts 6.4×10^7) obtained from an as-received single crystal of MgO(100). Part (b) of Figure 26 shows the anisotropic part of the 2D-ACAR distribution obtained from a Li-implanted sample after annealing at 950 K (total number of counts 1.0×10^7). This distribution was collected at 4 keV positron

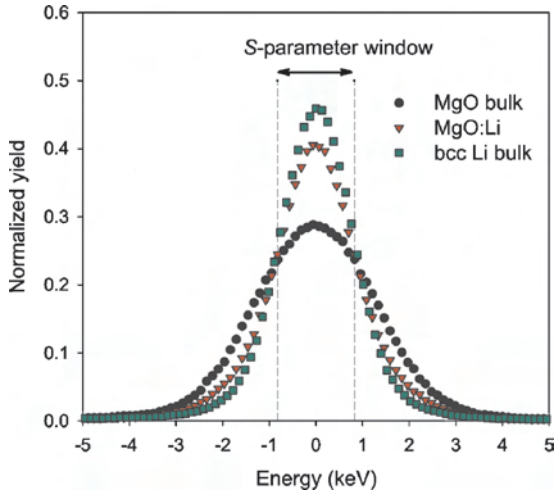


Figure 25. Experimentally obtained 511 keV Doppler broadening peaks for bulk MgO, bulk bcc Li, and MgO containing Li nanoclusters (after the 950 K annealing step).

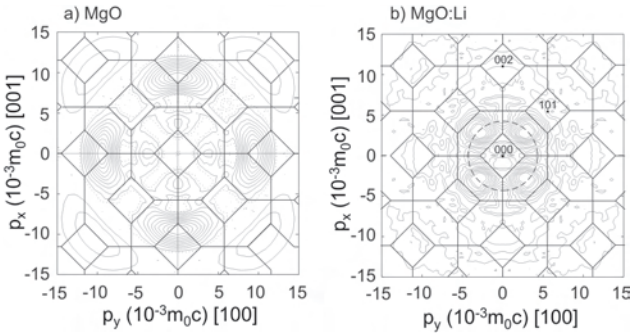


Figure 26. Anisotropic part of the [010]-axis projected momentum distribution obtained by 2D-ACAR measurements for (a) bulk MgO(100) and (b) MgO containing Li crystalline nanoclusters. The average Fermi cutoff for bulk Li of 4.3 mrad (see Section 4.2) is displayed as a dashed circle. The distributions have been symmetrized by reflection with respect to the (100) and (001) mirror planes. The contour interval is 10% (5%) of the peak height for solid (dotted) contours. The projection of the Brillouin zones is shown based on the lattice parameters of MgO (a) and fcc Li (b).

implantation energy, which corresponds to a mean positron implantation depth at the center of the lithium nanocluster layer (MgO:Li material). This energy was chosen on the basis of depth profiling performed on the 2D-ACAR sample by means of Doppler broadening PBA. Both distributions have been corrected for the difference in resolution between the directions parallel and perpendicular to the sample surface (1.1×1.4 mrad^2) and were symmetrized by reflection with respect to the (100) and (001) mirror planes. The anisotropy plot of the MgO:Li layer is strikingly different from the bulk MgO anisotropy plot. The bulk MgO anisotropy is still observable in Figure 26b, but another, very dominant anisotropic contribution is present in the central part of the figure. The major part of this contribution has a nearly perfect fourfold symmetry with peaks positioned near the Fermi radius of lithium. Therefore, we attribute the 2D-ACAR anisotropy in the center of Figure 26b to annihilation in bcc or fcc lithium clusters. Treilleux and Chassagne [112] found by means of electron diffraction and X-TEM analysis that, depending on the nanocluster size, metallic lithium nanoclusters in MgO can adopt either the fcc or a rotated bcc crystal structure. Whether the lithium clusters found in the 2D-ACAR distribution of Figure 26b are in the fcc or rotated bcc phase has not yet been established unequivocally. The fact that the anisotropic contributions displayed in Figures 26a and 13b both have the same symmetry and identical mirror planes shows that there is a distinct orientation relationship between the crystalline Li clusters and the MgO host matrix. Therefore, the lithium clusters will be coherent or semicoherent with the MgO crystal lattice. Comparing the ion implantation settings and annealing temperature used in this work with the experimental parameters of [112], the presence of fcc lithium is most likely. In that case, the lithium nanoclusters are in a simple cube-on-cube orientation relationship with the MgO host matrix. Cube-on-cube bcc Li is excluded by the anisotropy in Figure 26b as it would produce peaks along the $\langle 100 \rangle$ directions. The anisotropy in the central part of Figure 26b agrees very well with the results of preliminary calculations

performed on the momentum distribution of fcc Li using the KKR method. The free electron Fermi sphere for fcc Li ($a_0 = 4.4$ Å, cf. Table 5) is displayed in Figure 26b as a circle of radius 0.591 a.u., equivalent to 4.3 mrad. The peaks in the anisotropy then derive from bulges of the Fermi surface pointing toward the hexagonal Brillouin zone faces.

The fraction of positron annihilations in Li nanoclusters f_{Li} was also estimated from the 2D-ACAR distribution, using Eq. (36) in combination with a VEPFIT layer fraction analysis (giving the relative contributions of the layers to the annihilation distribution at a certain positron implantation energy). The fraction found (90%) was not significantly different from the one derived from the PBA results discussed in Section 4.4.1.

4.4.3. Positron Affinities

It was found that the fraction of annihilations in Li as deduced from the Doppler broadening PBA spectra is $92 \pm 5\%$ in the ion implanted layer. It will be shown that the effective positron confinement can be ascribed to the different positron affinities of MgO and Li. These positron affinities were calculated using the various models discussed in Section 4.2. The results of calculations of the positron lifetime, affinity, and electron and positron chemical potentials for Li are shown in Table 7. Very similar lifetimes and affinities are obtained for bcc and fcc Li. This is related to the fact that the atomic volumes of bcc and fcc Li are almost equal (Tab. 5). The affinities found are in good agreement with theoretical values from the literature [105, 125].

The electron and positron chemical potentials, positron affinity, and lifetime for MgO determined using three different models are given in Table 8. The IM and GGA approaches yield values for the bulk lifetime that are close to the experimentally obtained lifetime τ of 150–170 ps in MgO [126, 127], which is not the case for the SM results. We note that the incorporation of empty spheres into the MgO lattice (MgO II case) leads to a slight increase of the lifetime values because the interstitial space is better described compared to the case without empty spheres (MgO I). Table 8 also includes the calculated lifetimes by Puska et al. [115]. In the present work, the lifetimes found using the SM and the IM are somewhat higher and lower, respectively, than the values reported by Puska et al., while the same SM and IM models have been used. This is mainly due to a different treatment of the core electrons. In addition, the electron density of MgO in [105] was approximated by the superposition of atomic densities without consideration of the charge transfer between Mg and O atoms. Anyhow, it turns out that the semiconductor model does not describe reasonably the lifetime experiment (both in the present work and in [105]) so that we will not consider this model when calculating the difference between positron affinities of MgO and Li.

Table 7. Calculated positron lifetime, positron affinity, and electron and positron chemical potentials for bcc and fcc Li.

		LDA			GGA		
	μ_- (eV)	τ (ps)	μ_+ (eV)	A_+ (eV)	τ (ps)	μ_+ (eV)	A_+ (eV)
bcc Li	−2.8	299	−5.2	−8.0	283	−5.0	−7.8
fcc Li	−2.8	297	−5.3	−8.1	282	−5.0	−7.8

Table 8. Calculated lifetimes, positron affinities, and electron and positron chemical potentials for MgO without empty spheres (MgO I) and including empty spheres (MgO II).

	GGA				SM			IM		
	μ_- (eV)	τ (ps)	μ_+ (eV)	A_+ (eV)	τ (ps)	μ_+ (eV)	A_+ (eV)	τ (ps)	μ_+ (eV)	A_+ (eV)
MgO I	−2.9	139	−3.2	−6.1	121	−4.0	−6.9	141	−2.6	−5.5
MgO II	−1.0	140	−5.0	−6.0	126	−5.5	−6.5	147	−4.0	−5.0
MgO[a]					119			167		

[a] From [115].

The values of A+ presented in Table 8 depend obviously on the models used to calculate them. The lowest value comes from the SM, whereas the IM yields the largest value. Also, the affinity results depend slightly on whether the empty spheres are included or not. In the following considerations we will employ the MgO II case. The positronic energy step at the MgO//Li interface can now be calculated using Eq. (39). Combining the GGA and IM positron affinities for MgO II (Tab. 8) and the GGA results for bcc and fcc lithium (Tab. 7; LDA and GGA results differ only marginally), the positronic potential drop at the MgO//Li interface equals 1.8 eV using the GGA and 2.8 eV using the IM. Regardless of this uncertainty in the affinity difference, it is plausible to assume that clusters of Li in MgO act as positron traps.

In this context it is useful to mention recent work [118] in which the positron affinity was studied in several polytypes of SiC, which is a semiconductor. It was found that all theoretical approaches (GGA, SM, IM) overestimate (in magnitude) the positron affinity in SiC. This was ascribed to certain inadequacies in the theoretical description of both the electron and the positron part of the problem. Considering this conclusion and the aforementioned experimental values of the positron affinity for MgO [108, 109] it turns out that the values of A_+ calculated here for MgO are probably too large (in magnitude). Further theoretical and experimental work is therefore needed to clarify this point. A final remark in this section concerns the sensitivity of the values of the positron affinity with respect to the choice of the LDA exchange-correlation potential for electrons. In addition to the exchange-correlation potential of Von Barth and Hedin [117] we also examined the potential of Ceperley and Alder [128] as parameterized by Perdew and Zunger [129]. Surprisingly, the calculated positron affinities increase (magnitudes decrease) by about 0.5 eV in all cases presented in Table 7. The inspection of electron and positron chemical potentials leads to the conclusion that this effect can be attributed to the change of the electron chemical potentials only. It shows that, although the resulting difference of the positron affinity of MgO and Li remains unchanged, one has to be careful when comparing affinity results obtained using different exchange-correlation potentials for electrons (cf. [130]).

4.5. Discussion

This section shows a very high fraction of annihilations in lithium, indicating a very effective positron confinement in Li nanoclusters. Two aspects are important when considering the efficiency of positron confinement in nanoclusters: (a) the probability of trapping and reflection when a positron encounters a nanocluster acting as a potential well, and (b) the probability of reaching a nanocluster, assuming that the positron is thermalized in the embedding MgO material.

The kinetic energy of delocalized positrons is of the order of tens of meV, while the size of the potential step at the MgO//Li interface is a few eV (1.8 eV using the GGA and 2.8 eV using the insulator model). Thus, the depth of the potential well as determined by the difference in positron affinities is two orders of magnitude larger than the kinetic energy of the positron after thermalization. Therefore, when a positron is thermalized and diffuses within the embedding MgO, the probability of reflection when encountering the MgO//Li interface is negligible at room temperature. Every positron reaching a lithium nanocluster will thus be confined in that cluster. Once inside the cluster, the positron will be reflected at the walls. A necessary condition for the process described is that the size of the nanoclusters is large enough to accommodate at least one bound state. Applying the three-dimensional spherical well model with a finite potential, the critical radius is given by Eq. (40). Substituting a value of 1.8 eV for the depth of the potential well then yields a critical cluster diameter of 0.4 nm. As the confinement of positrons is quite obvious from the results described in Sections 4.4.1 and 4.4.2, the actual size of the nanoclusters will be larger than this value.

The probability of encountering a nanocluster is determined by the positron diffusion length and the concentration and spatial distribution of the Li clusters. Trapping of the vast majority of the positrons is only possible if the mean distance between the Li nanoclusters is smaller than the diffusion length for positrons diffusing in MgO within the MgO:Li layer. Here it is assumed that the positron is present in MgO after thermalization, as Li occupies only 3 vol% of the MgO:Li layer. In the diffusion-limited trapping model, described by Hautojärvi and Corbel [131], the diffusion to the defect is slow compared to the transition to the localized state. The depth of the nanocluster potential well is of the order of a few eV, so this condition is well satisfied. The possibility of detrapping is neglected. Assuming spherically shaped clusters, the fraction of positrons annihilating in the clusters is given by

$$f_{\mathrm{Li}} = \frac{\kappa}{\kappa + \lambda_{\mathrm{MgO}}} = \frac{4\pi r_{\mathrm{cl}} D_{+} c_{\mathrm{cl}}}{4\pi r_{\mathrm{cl}} D_{+} c_{\mathrm{cl}} + \lambda_{\mathrm{MgO}}} \tag{44}$$

where κ is the positron trapping rate with respect to the clusters (s^{-1}), λ_{MgO} is the annihilation rate in MgO (s^{-1}), r_{cl} is the radius of the cluster (m), c_{cl} is the concentration of clusters (m^{-3}), and D_{+} is the diffusion coefficient ($m^2\ s^{-1}$). The diffusion coefficient is related to the diffusion length L_{+} (m) and to the positron lifetime τ (s) as

$$L_{+} = \sqrt{D_{+}\tau} \tag{45}$$

By means of Eqs. (44) and (45), the fraction of annihilations in lithium clusters can be estimated. No detailed information is available on the size of the nanoclusters, so we perform an order-of-magnitude calculation. The mass density of bulk Li is 543 kg m^{-3}, so the 10^{16} implanted Li atoms occupy a total volume of 2.1×10^{-13} m^{-3} independent of the cluster size. If the radius of the lithium nanoclusters is, for example, 5.0 nm [112], the total number of clusters is 4.1×10^{11} and the concentration of clusters will be 5.8×10^{22} m^{-3} considering that all clusters are present in the ion implantation layer with a thickness of 70 nm (estimated from VEPFIT, Tab. 6) and an area of 1.0×1.0 cm^2. At this concentration, the average intercluster distance is 26 nm. The annihilation rate in MgO equals 6.5×10^9 s^{-1} as it is the inverse of the lifetime τ in MgO, 155 ps [126]. A value for the diffusion length for positrons in MgO of 50 nm can be assumed (equal to the diffusion lengths for layers 1 and 3 in the VEPFIT model). Substituting Eq. (45) into Eq. (44) and using the numerical values given previously, the fraction of annihilations in Li, f_{Li}, can be calculated. This yields a fraction of 90%, which agrees very well with the experimental results. Although the cluster size and the diffusion length are estimated, the calculation demonstrates that the annihilation fraction in lithium can be lower than 100% due to diffusion-related processes even if every positron encountering a nanocluster is trapped in that cluster. The calculated fraction is strongly dependent on the cluster size and the diffusion length. For clusters with a radius of 3–7 nm and a diffusion length of 50–100 nm, the trapped fraction varies from 82% to 99%.

Finally, the possibility should be discussed that positrons are not present in either MgO or Li but are trapped at the interface, corresponding to the situation depicted in Figure 21c. This phenomenon is observed, for example, for the Si//SiO_2 interface [132]. Open volume defects or lattice mismatch at the interface can induce positron trapping. No evidence of positronium formation was found in the analysis of the 2D-ACAR distribution, showing that large vacancy-type defects are not present. Positrons can also trap in interface defects that are too small for the formation of positronium. However, positron annihilation distributions corresponding to positron annihilation at interfaces usually bear characteristics of both materials. Thus, if positron trapping would occur at the MgO//Li interface, a considerably higher contribution from MgO to the Doppler broadening and 2D-ACAR distributions would be expected. Therefore, interface trapping in the case of the MgO//Li interface is not likely on the basis of these observations. On the other hand, the fraction of annihilations in Li nanoclusters might be influenced by local electric fields near the interface. These can be present as a result of charge redistribution associated with the equalization of the Fermi levels of MgO and Li in thermal equilibrium.

4.6. Conclusions

In this section it has been shown that positrons are very effectively confined in crystalline lithium nanoclusters that are embedded in MgO. The fraction of positron annihilations in lithium nanoclusters has been deduced by analysis of the Doppler broadening PBA spectra and found to

be $92 \pm 5\%$. This annihilation fraction is very high, considering the fact that the lithium content in the MgO:Li layer is only 1.3 at%. The anisotropic contribution of the 2D-ACAR distribution collected in the layer containing Li nanoclusters shows a fourfold symmetry (most likely indicating an fcc crystal structure) with peaks at positions near the average Fermi radius of lithium. This proves that crystalline bulk lithium is present. The symmetry of the 2D-ACAR anisotropic contributions also shows that the lithium clusters are coherent or semicoherent with the MgO host matrix. The difference in positron affinities of lithium and MgO yields a positronic potential step down at the MgO//Li interface. Therefore, the nanoclusters act as a potential well for positrons with a depth equal to the difference in the positron affinities. These positron affinities were calculated using the LMTO-ASA method, and values for the depth of the positronic potential well of 1.8 and 2.8 eV were obtained using the GGA and the insulator model, respectively. The depth of the potential well is so large that the trapping probability of positrons meeting a nanocluster is approximately equal to 1.

5. FORMATION OF SOLID Kr NANOCLUSTERS IN MgO

5.1. Introduction

Ion implantation of noble gas atoms in matrices offers the opportunity to create nanoclusters of van der Waals solids at room temperature, which is convenient for studying the optical and structural properties of these exotic materials. Here it should be realized that the material properties of small clusters (i.e., with sizes of a few nm or less) usually deviate from that of the bulk material. Apart from the fundamental interest in the material properties, a technological application of Kr implantation in MgO is the fabrication of ultrathin MgO slabs with a thickness of 100–200 nm [133]. Solid Kr clusters created by means of ion implantation were already observed in the metals Ti, Ni, Cu, Mo, and Au [134, 135] in graphite and Grafoil [136] and once in MgO [137]. The advantage of embedding the clusters in transparent materials such as MgO is that it facilitates optical studies. Most of the work mentioned concerns the analysis of diffraction patterns in TEM and X-ray absorption studies. Here, we discuss solid Kr clusters in the TEM in high-resolution mode, similar to the high-resolution work already performed on solid Xe clusters [138].

Apart from TEM, we also used optical absorption spectroscopy and Doppler broadening PBA [8] to follow the defect evolution in the sample. If positrons are effectively trapped in Kr nanoclusters, this allows investigation of the electronic structure of solid Kr by means of positron annihilation techniques. If the positron wave function is spatially confined within the nanoclusters in three dimensions, we speak of positron confinement [99, 139–141]. Kr is an important candidate for studying the effect of positron confinement because of the very low positron affinity of Kr [142], which means that the nanoclusters are very likely to act as a three-dimensional potential well to positrons. Positron trapping in solid Kr nanoclusters embedded in metals (Cu, Ni) was studied previously using various positron annihilation techniques [143, 144]. The main conclusion from these investigations is that positrons trap at the Kr–metal interface.

5.2. Experiments

Several epi-polished MgO(100) single crystals of size $10 \times 10 \times 1$ mm^3 were implanted at room temperature with 280 keV Kr ions to a dose of 3×10^{16} ions cm^{-2} and subsequently annealed in ambient air for periods of 30 min at temperatures of 900 and 1100 K, respectively. After ion implantation and after each annealing step, the evolution of ion implantation defects was monitored with optical absorption spectroscopy and Doppler broadening positron beam analysis using a monoenergetic positron beam with a variable acceleration energy of 0–30 keV. X-TEM was applied after the 1100 K annealing step. The TEM was performed using a JEOL 4000 EX/II operating at 400 kV (point-to-point resolution 0.165 nm; see Section 2). The specimen preparation is discussed elsewhere [79]. The sample treatment and main observations are listed in Table 9.

5.3. Results

First, results of optical absorption spectroscopy will be presented. Next, the X-TEM results will be shown, focusing both on the solid Kr nanoclusters and on the defects created in the MgO. Finally the result of positron beam analysis is shown as an additional method to obtain depth-resolved information on the defect evolution in the composite material during the sample treatment. Possible trapping in Kr nanoclusters is discussed in terms of the positron affinity.

5.3.1. Optical Absorption Spectroscopy

The results of the optical absorption spectroscopy are displayed in Figure 27. After ion implantation, two types of defects can be distinguished. At a wavelength of 570 nm, there is an absorption peak that can be attributed to V-centers (Mg vacancies) and at a wavelength of 250 nm; there is another absorption peak that can be attributed to F-centers (O vacancies) [145]. Both absorption peaks vanish after annealing at 900 K.

Table 9. Sample treatment and main experimental observations.

Sample treatment	
Ion implantation	3×10^{16} Kr ions cm^{-2} at an energy of 280 keV
Thermal anneal	at 900 K and 1100 K for a period of 30 min
Results	
Optical absorption	F- and V-centers present after implantation; dissociation after 900 K anneal.
X-TEM	Cubical, solid Kr clusters at 70–130 nm depth, cluster size 2–4 nm. Cubical nanovoids at 15–30 nm depth, cluster size 2–5 nm.
PBA	Defect agglomeration during annealing. High S parameter in nanovoids layer, S parameter in Kr cluster layer higher than that of bulk MgO and lower than the S parameter of MgO with defects.

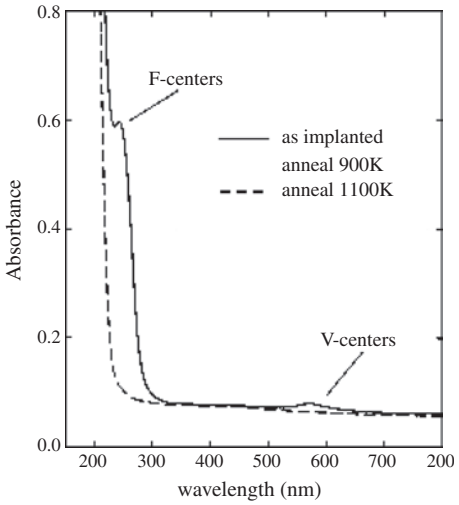


Figure 27. Optical absorption spectra of an MgO(100) crystal implanted with 280 keV 3×10^{16} Kr ions cm^{-2}, as implanted and after annealing in air at various temperatures.

5.3.2. Cross-Sectional TEM

X-TEM has been performed on a sample after the 1100 K anneal step. Figure 28 shows an overview image. The area between 70 and 130 nm depths is a subsurface layer that contains Kr nanoclusters. The rectangular area in Figure 28 (marked with the white brackets) is enlarged in the HRTEM image of Figure 29. Here the Kr nanoclusters can be clearly observed. The clusters are rectangularly shaped and the sizes vary from 2 to 5 nm. Moiré fringes are present due to the lattice mismatch between solid Kr and MgO. The morphology of the nanoclusters is determined by the interface energy corresponding to the various MgO//Kr interface facets; the

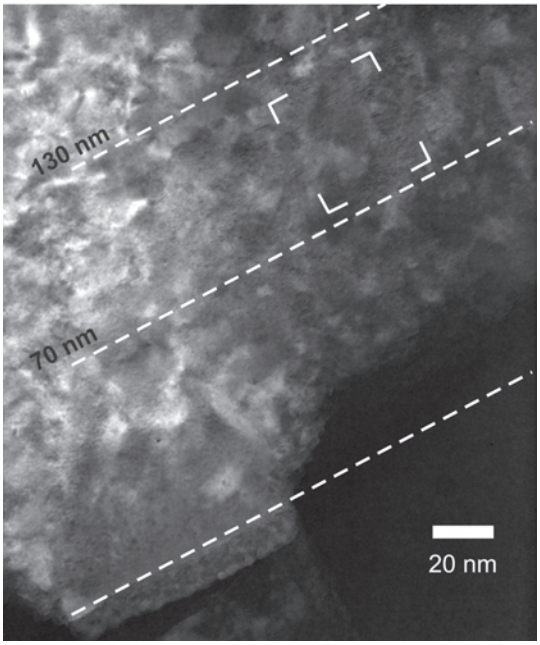


Figure 28. TEM overview image of Kr-implanted MgO. Solid Kr nanoclusters are observed in a band at a depth of 70–130 nm.

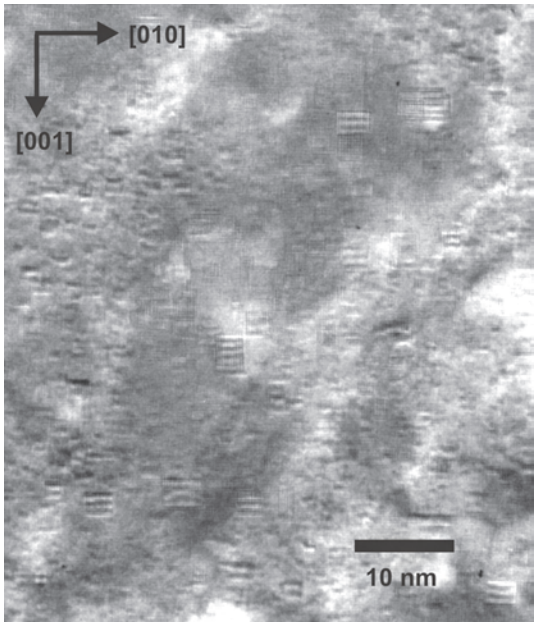


Figure 29. Solid Kr nanoclusters at a depth of 75–120 nm with Moiré fringes caused by the lattice mismatch between solid Kr and MgO.

shape of the cluster can be constructed using the Wulff diagram [146, 147] if the interface energies are known. In this particular case, it is clear that the {100} interface is energetically favored to the {110} and {111} interfaces. The interface energy equals the sum of the surface energies of MgO and Kr minus the work of adhesion:

$$\gamma_{\mathrm{MgO||Kr}} = \gamma_{\mathrm{MgO}} + \gamma_{\mathrm{Kr}} - W_{\mathrm{ad}} \tag{46}$$

Considering that the van der Waals bonding of the noble gas Kr atoms is very weak, the interface energy of Kr and the work of adhesion will be negligible with respect to the surface energy of MgO (i.e., by approximation the MgO//Kr interface energy equals the MgO surface energy). Watson et al. [148] calculated surface energies of 1.25 J m^{-2} for MgO {100}, 1.87 J m^{-2} for MgO{110}, and 2.5 J m^{-2} for MgO{111}. The last two values are valid for microfaceted surfaces, but the surface energies for unfaceted {110} and {111} MgO surfaces are even higher [148]. Considering the Wulff diagram the absence of the {110} facets in the rectangular Kr nanoclusters means that the MgO{110} surface energy should be at least $\sqrt{2}$ times larger than the MgO{100} surface energy, which is in agreement with the theoretical values given.

Figure 30 shows a magnification of the large cluster present in the center of Figure 29. There are 5 moiré fringes in 18 MgO lattice spacings. Because Kr exhibits an fcc structure in fcc host matrices [135] and the Kr clusters are in a cube-on-cube orientation relationship with the MgO, the following relationship can be used to calculate the lattice parameter of solid Kr, a_{Kr}

$$\frac{1}{d_{\mathrm{fringes}}} = \left| \frac{1}{d_{\mathrm{MgO}}} - \frac{1}{d_{\mathrm{Kr}}} \right| \tag{47}$$

With a lattice parameter a_{MgO} of 4.212 Å ($d_{\mathrm{MgO,\,200}} = 2.106$ Å), this yields $d_{\mathrm{Kr,\,200}} = 18/(18-5) \cdot d_{\mathrm{MgO,\,200}} = 2.9$ Å

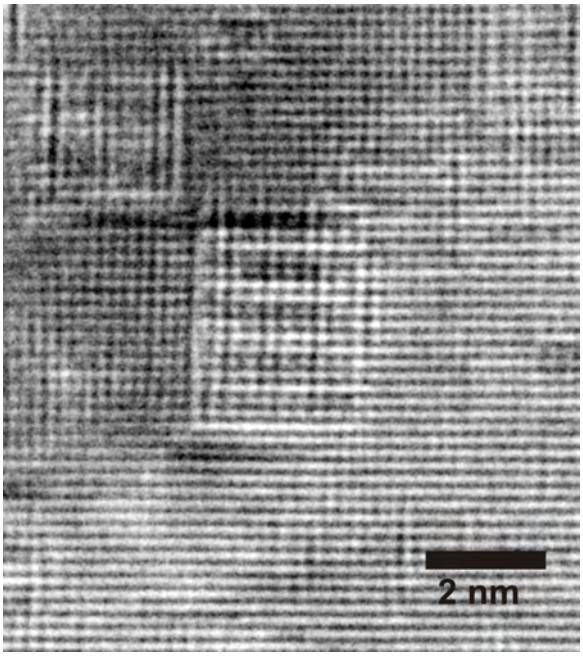


Figure 30. High-resolution TEM image of a solid Kr nanocluster (also present in the center of Figure 29). The lattice parameter of the solid Kr can be derived from the Moiré fringes (see text).

and thus the lattice parameter for solid Kr, a_{Kr}, equals 5.8 Å. This corresponds to the lattice parameter of frozen Kr at a temperature of 110 K [149]. In a similar fashion, the Kr lattice parameters in the other clusters present in Figure 29 can be analyzed and values of 5.3–5.8 Å are found. These values for a_{Kr} correspond reasonably well with the values of 5.0–5.5 Å reported in the literature [134, 137]. The important advantage of the use of moiré fringes in comparison with the experimental methods mentioned in the Introduction is that the lattice parameter can be determined for every Kr cluster separately. In Figure 31, the Kr lattice parameter is plotted as a function of the cluster size. It is clear that there is a positive correlation between a_{Kr} and the cluster size; the

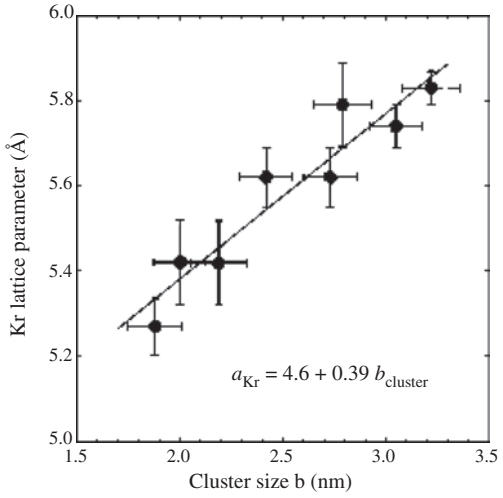


Figure 31. Relationship between the Kr lattice parameter and the cluster size.

straight line plotted in the figure is a linear least squares fit to the data. The cluster size (defined as the cube root of the nanocluster volume) cannot be determined accurately because the cluster dimension perpendicular to the plane in Figure 29 or 30 cannot be determined. Therefore, the cluster size was estimated as the square root of the cluster area that is visible in the figures. The lattice parameter in small clusters is smaller due to the higher pressure in small clusters. The lattice parameter can be related to the pressure using the Ronchi equation of state (EOS). A lattice parameter of 5.3–5.8 Å for fcc Kr corresponds to a molar volume of 22–30 cm^3. Using the EOS for solid Kr [150] at a temperature of 300 K, the pressure in the largest cluster is 0.6 GPa and the pressure in the smallest cluster is 2.5 GPa. These pressures are in reasonable agreement with the minimum pressure of 0.83 GPa that is necessary for the formation of solid Kr [137]. It is also in good agreement with the following equations, describing the relationship between pressure and interface energy. The energy of the cluster system is minimized when

$$dE = p\,dV - \gamma\,dS = 0 \quad \Rightarrow \quad \frac{dE}{da} = p\frac{dV}{da} - \gamma\frac{dS}{da} = 0 \tag{48}$$

For the case of a cube with size b, volume $V = b^3$, and surface $S = 6b^2$, it follows that

$$p = \frac{4\gamma}{b} \tag{49}$$

Since the interface energy is completely determined by the surface energy of MgO [see Eq. (46) and the previous discussion], the pressure in the cluster can be calculated. It should be realized that Eqs. (48) and (49) are only valid when the material is in thermal equilibrium (i.e., at the temperature at which the clusters were formed). The pressure and surface energy mentioned in Eq. (49) should therefore be considered at 1100 K. The MgO{100} surface energy at 1100 K is 1.17 J m^{-2} [151]. Using Eq. (49), a cluster size of 3 nm then yields a pressure of 1.7 GPa at 1100 K. How can this pressure be compared with the pressures of 0.6–2.5 GPa found in the electron microscope at 300 K? MgO is a rather incompressible material with a linear thermal expansion coefficient that varies with temperature. Over the temperature interval of 300–1100 K, the average linear thermal expansion coefficient a_{lin} is 12×10^{-6} K^{-1} [152]. Since MgO is very incompressible compared to Kr, we assume that the shrinkage of the cluster is completely determined by the MgO. With $a_{vol} = 3a_{lin}$, the change in volume of the cluster is 2.9%. Therefore, the molar volumes of 22–30 cm^3 as found at room temperature correspond to molar volumes of 23–31 cm^3 at 1100 K. Using again the Ronchi equation of state [150] but now for a temperature of 1100 K, we find pressures of 1.5–4.6 GPa. This is in good agreement with the pressure of 1.7 GPa that was found from Eq. (49).

Figure 32 shows an enlarged TEM image of the subsurface area down to a depth of 60 nm. It is clear that a band of rectangularly shaped nanovoids is present at a depth of 15–30 nm. These large vacancy clusters are 2–5 nm in size and are rectangularly shaped, similar to the nanovoids found previously in ion implanted MgO [153]. The rectangular shape is again caused by the different surface energies of

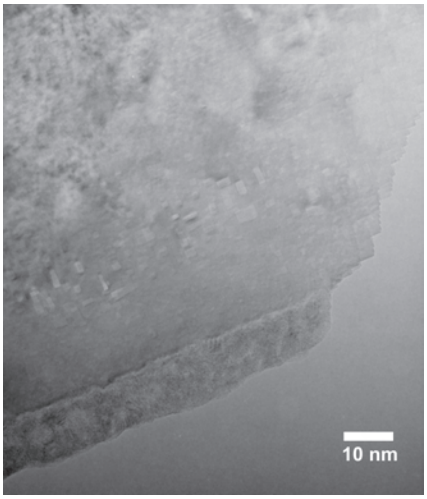


Figure 32. TEM image showing the vacancy cluster band located at a depth of 15–30 nm.

the various MgO facets (see the discussion on the rectangular shape of the Kr clusters). These voids are formed due to accumulation of vacancies created by the Kr implantation. It is, however, peculiar that this vacancy band is located so far away from the main Kr cluster band, because normally the damage depth distribution and the Kr depth distribution from the same implantation are overlapping to a large extent.

Figure 33 shows the Kr depth distribution and the damage created by ion implantation as simulated by means of the SRIM2000 code [83]. Here displacement energies of 55 eV were used for both Mg and O [154]. For the applied dose of 3×10^{16} ions cm^{-2}, the peak value of the Kr concentration is 4.1 at% and the peak damage level is 37 dpa (displacements per target atom). First of all, it is clear that the predicted Kr depth distribution is in excellent agreement with

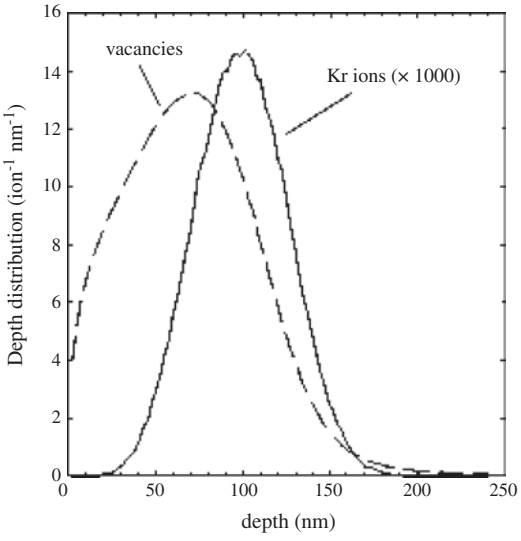


Figure 33. Kr depth distribution (after implantation, prior to thermal anneal) and damage depth distribution as calculated using the SRIM2000 code.

the location of the Kr nanocluster layer at a depth of 70–130 nm as observed by TEM in Figure 28. It is also clear that the predicted damage depth distribution is located shallower than the predicted Kr distribution. However, this does not explain the presence of a distinct vacancy cluster band at a depth of 15–30 nm as observed by TEM: in the SRIM simulation the vacancy and Kr distribution are not separated but overlap to a large extent. A possible explanation is that the vacancies and Kr atoms recombine at the intermediate depth range of 30–70 nm: the presence of Kr in the vacancy clusters slows down the mobility of Kr-filled vacancy clusters so that large vacancy clusters are not formed. At the same time, the concentration of Kr in this intermediate layer is not high enough to form solid Kr nanoclusters that are big enough to be observed by means of TEM.

5.3.3. Positron Beam Analysis

The defect evolution during the annealing procedure was monitored by Doppler broadening PBA [8]. Annihilation of positrons with electrons in solids yields information on the momentum distribution of these electrons. The electronic momentum distribution is reflected in the Doppler broadening of the 511 keV annihilation peak. Positron annihilation with low-momentum valence or conduction electrons results in a small Doppler shift, contributing to the center of the peak. Annihilation with high-momentum core electrons results in a large Doppler shift, contributing to the wings of the 511 keV annihilation peak. The shape of the 511 keV peak is characterized by the so-called *S* (shape) parameter, which gives the ratio of the number of counts in the center of the peak to the number of counts in the whole peak (see Section 2). By using a monoenergetic positron beam with variable energy, the *S* parameter can be recorded as a function of depth. However, at larger depths the depth resolution decreases due to broadening of the implantation profile of the implanted positrons.

The *S* parameter is displayed in Figure 34 as a function of positron implantation energy. The average positron implantation depth is indicated on top of the figure. The solid lines show the results of the VEPFIT simulation that will be discussed. It is clear that after implantation, the *S* parameter increases with respect to reference MgO in the ion implantation zone and it increases further after annealing at 900 and 1100 K over a wide range. The *S* parameter in ion-implanted MgO always increases after ion implantation due to the efficient formation of vacancy-type defects in the Schottky material MgO. It is observed that the *S* parameter increases after annealing at 900 K. Simultaneously, the optical F- and V-centers dissociate after annealing at this temperature as observed with optical absorption spectroscopy (see Fig. 27). This clearly indicates aggregation of vacancy-type defects (the larger the vacancy clusters, the larger the *S* parameter). The peak at 1.5 keV positron implantation energy in Figure 34 can be attributed to the presence of the vacancy cluster band that was observed with TEM in Figure 32. These voids are large enough to form Ps, a hydrogenlike bound state of a positron (e^+) and an electron (e^-). Positronium formation leads to a dramatic increase in the *S* parameter. In order to analyze the PBA results in more detail, the VEPFIT code [49] was used to find the

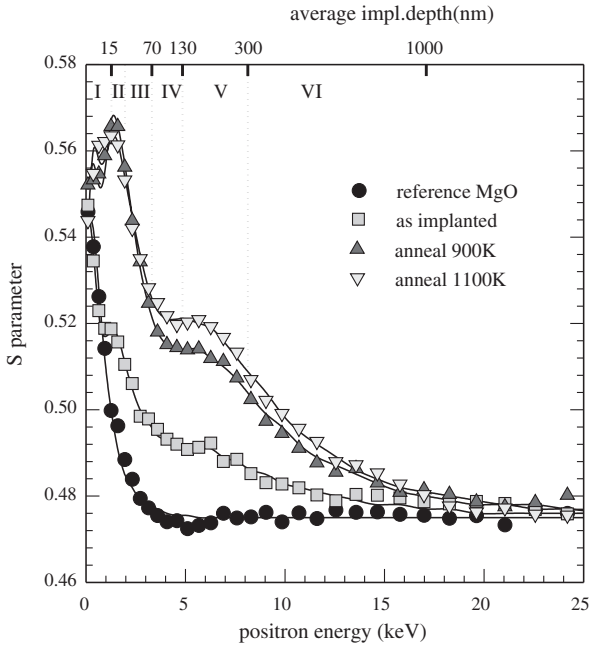


Figure 34. *S* parameter vs positron energy measured for reference MgO and for MgO:Kr after implantation and after thermal annealing treatments. The scale at the top of the graph indicates the average positron implantation depth. The Roman numerals are the layer numbers (see Table 10) and the solid lines are the result of VEPFIT fitting (see text).

S parameter corresponding to every defect layer. The *S* parameter found at a certain positron energy in Figure 34 consists of contributions of various layers due to the broadness of the positron implantation profile (which is increasing with increasing positron energy) and due to positron diffusion processes. The VEPFIT code takes these effects into account and yields the depth-resolved *S* parameters for every defect layer. Based on the TEM observations, a six-layer model was used as defined in Table 10. The chosen diffusion lengths are reasonable estimates. When considering the number of clusters in Figures 29 and 32, it is clear that here the diffusion length will be 5 nm or less. The layers and the corresponding depths are indicated at the top of the graph in Figure 34. Layer V with a depth of 130–300 nm is the so-called ion implantation tail. Due to channeling effects, the implantation damage in MgO extends much further than the maximum range as predicted by SRIM (Fig. 33) as this code assumes the atoms in the target material to be randomly distributed. Only for the *S* parameter curve of reference (nonimplanted) MgO, a single-layer model was used with the *S* parameter value of bulk MgO, 0.475.

Table 10. Model used for VEPFIT simulation.

Layer	Description	Depth (nm)	*S* parameter	Diff. length (nm)
I	MgO with defects	0–15	S_{def} (fitted)	5
II	nanovoids	15–30	S_{voids} (fitted)	5
III	MgO with defects	30–70	S_{def} (fitted)	5
IV	Kr clusters	70–130	S_{Kr} (fitted)	5
V	MgO with defects	130–300	S_{def} (fitted)	20
VI	MgO bulk	>300	$S_{MgO} = 0.475$	100

Note: The *S* parameters of layers I, III, and V are fitted with the same parameter.

The solid lines in Figure 34 are the result of VEPFIT fitting and it is clear that the model as discussed is well able to fit the experimental *S* parameter curves. Figure 35 shows the depth-resolved *S* parameters found using VEPFIT for the sample after the 1100 K annealing step (where X-TEM was also performed). As expected, the *S* parameter in the nanovoids layer (layer II) is rather high, 0.59. The *S* parameter in the layer with Kr nanoclusters (layer IV) is lower than that of the adjacent layers but higher than that of bulk MgO. This explains the "dip" in the *S* parameter curve in Figure 34 at 4.5 keV positron implantation energy. The fact that the *S* parameter in the layer with Kr nanoclusters is not much different from adjacent layers points to little interaction of positrons with Kr nanoclusters. The reason that the *S* parameter in the nanocluster layer is lower than in the adjacent layers (Figs. 34 and 35) is probably due to Kr filling of vacancy-type defects. Thus there are less open volume defects for the positrons to trap in and the value of the *S* parameter is closer to that of bulk MgO.

In order to investigate the behavior of positrons with respect to the nanoclusters in more detail, a 2D-ACAR (two-dimensional angular correlation of annihilation radiation) experiment was carried out on the nanoclusters after the 1100 K annealing step. 2D-ACAR allows a detailed investigation of the electronic structure of a material [155]. Experimental details of the setup can be found in Section 2 and [140], where positron confinement in Li nanoclusters in MgO is investigated. The positrons were implanted at 4.5 keV energy, corresponding to implantation in the center of the Kr nanocluster layer. No anisotropic contribution other that that of MgO could be observed in the 2D-ACAR spectrum after accumulation of 7×10^6 2γ-coincidence events [156] thus supporting the idea of negligible trapping

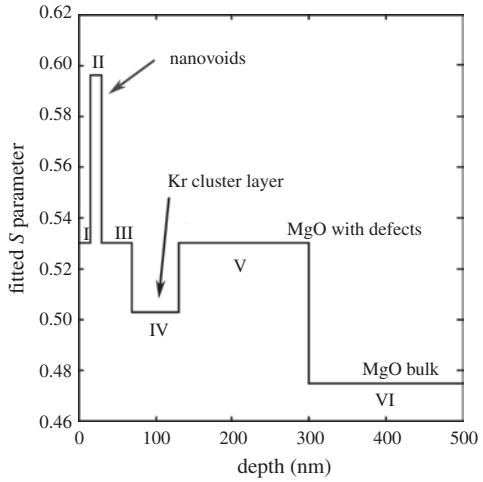


Figure 35. Depth-resolved *S* parameter as a function of depth in the sample after the 1100 K annealing step, as found using the VEPFIT code. The applied six-layer model is given in Table 10. The layers are indicated in Roman numerals.

in and interaction with Kr clusters. This is in contradiction with the expectation that positrons are very effectively trapped in Kr nanoclusters because of the very low positron affinity of solid Kr [142] compared to that of MgO [121]. So why do the Doppler broadening and 2D-ACAR results not give evidence of positron trapping in the Kr nanoclusters?

5.3.4. Positron Trapping and Positron Affinities

Whether positrons trap in nanoclusters is dependent on the positron affinity. First we deduce the definition of positron affinity for metals, following the derivation by Puska et al. [157]. Let us first consider Figure 36. Both the electronic energy levels and the positronic energy levels in the same material are defined with respect to the crystal zero (CZ). The crystal zero level is defined as the Coulomb potential far away from a single atomic sphere [158]. The electron chemical potential μ_- is defined as the energy difference between the top of the conduction band (coinciding with the Fermi level) and the crystal zero. The positron chemical potential μ_+ is defined as the energy difference between the bottom of the positron bands $E_{+,0}$ and the crystal zero. Both μ_- and μ_+ are usually negative. The electronic work function φ_- (positron work function φ_+) is the work to be done to bring an electron (positron) from the Fermi level (positron ground level) to vacuum. The surface dipole potential step Δ is the potential difference for electrons between the vacuum level and the crystal zero. Positrons experience the same potential step, but then with opposite sign: $-\Delta$. The following relationships hold:

$$\varphi_- = -\mu_- + \Delta \qquad \text{and} \qquad \varphi_+ = -\mu_+ - \Delta \tag{50}$$

When two metal surfaces are in contact, electrons will flow from one material to the other until the Fermi levels are aligned, thereby establishing thermal equilibrium. This situation is shown in Figure 37. The result is an interface dipole with potential difference $\Delta_{\text{align}}^{\text{METAL}} = \mu_-^A - \mu_-^B$. For positrons, this potential step is equal in magnitude but opposite in direction. The difference between the lowest positron energies of material A and B equals [157]

$$\Delta E_{AB,+}^{\text{METAL}} = E_{0,+}^A - E_{0,+}^B = \mu_+^A - \mu_+^B + \mu_-^A - \mu_-^B \tag{51}$$

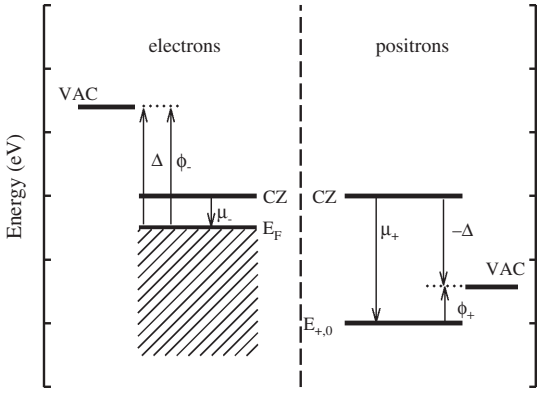


Figure 36. Schematic diagram showing the relevant electronic and positronic energy levels with respect to the CZ for the case of a metal.

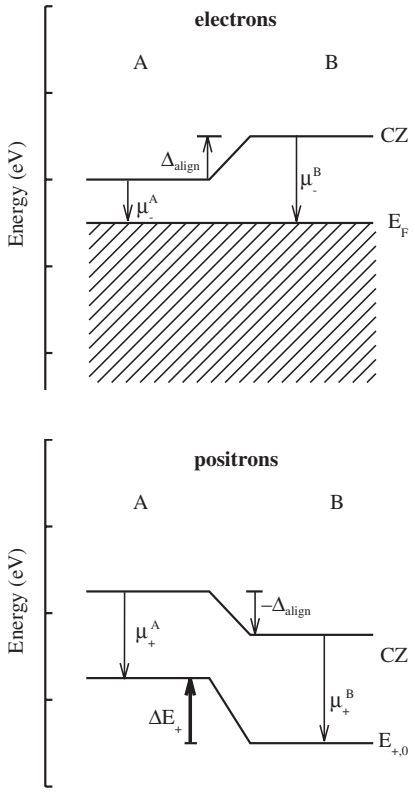


Figure 37. Schematic diagram showing alignment of the Fermi levels for the case of two metals in contact. Note that the alignment shift of the crystal zeros is $\Delta = \mu_-^A - \mu_-^A$ for electrons and $-\Delta$ for positrons.

which is also clear from Figure 37. When the positron affinity is defined as

$$A_+^{\text{METAL}} = \mu_+ + \mu_- = -(\varphi_+ + \varphi_-) \tag{52}$$

it follows from Eq. (51) that the difference in positron ground levels between the two metals is simply the difference between the positron affinities,

$$\Delta E_+^{AB} = A_+^A - A_+^B \tag{53}$$

Values for the positron affinity are commonly expressed in eV and are negative for most materials. The positron affinity is an entity that is characteristic of a material and it has been calculated for almost all metals [157], a number of semiconductors [159], a few insulators [160], and rare gases [94]. A necessary condition of the positron quantum confinement in a cluster (A) embedded in a host (B) is that $\Delta E_+^{AB} < 0$. Then the cluster represents a three-dimensional potential well for positrons. If $\Delta E_+^{AB} > 0$, the cluster acts as a potential barrier. Finally, the possibility should be considered that positrons trap at the interface.

The question now arises whether the concept of positron affinity as defined in Eq. (52) can be applied to an insulator–insulator interface. The alignment of the Fermi levels is expressed in the term $(\mu_-^A - \mu_-^B) = \Delta_{\text{align}}^{\text{METAL}}$ on the right-hand side of Eq. (51). This is true for metals where the position of the electronic chemical potentials coincides with the position of the Fermi levels. This is not true, however,

for insulators. In insulators the electronic chemical potential is defined as the distance between the top of the valence band and the crystal zero as depicted in Figure 38. E_V is the top of the valence band, E_C is the bottom of the conduction band, and E_F is the Fermi level positioned in the middle of the bandgap E_g. Figure 38 shows the electronic energy levels for MgO and Kr in contact for the case that the Fermi levels are aligned. It follows from the figure that $\Delta^{\mathrm{INS}}_{\mathrm{align}} = (\mu^A_- + \frac{1}{2}E^A_g - \mu^B_- - \frac{1}{2}E^B_g)$. The energies and energy levels displayed in the figure are to scale. The numerical values of the electronic and positronic entities for MgO and Kr are listed in Table 11 and are mainly obtained from [140, 142]. Figure 39 shows the corresponding positronic potential diagram for MgO and Kr in contact when the Fermi levels equalize. Considering Figures 38 and 39, it is clear that the difference between the positron ground potentials of the two materials is not described by Eq. (51) but by the following formula:

$$\Delta E^{\mathrm{INS}}_{AB,+} = E^A_+ - E^B_+ = \mu^A_+ - \mu^B_+ + \mu^A_- - \mu^B_- + \frac{1}{2}E^A_g - \frac{1}{2}E^B_g \quad (54)$$

Therefore, the positron affinity for insulators can be defined alternatively as

$$A^{\mathrm{INS}}_+ = \mu_+ + \mu_- + \frac{1}{2}E_g \quad (55)$$

so that Eq. (53) is still valid. Using Eq. (54) and the values gives in Table 11, we find a difference in positron ground potentials $\Delta E^{\mathrm{INS}}_{AB,+}$ of −6.7 eV. So $\Delta E_{AB,+} < 0$ and the Kr nanoclusters should act as potential wells, in contradiction with our experimental observations that do not show evidence of positron trapping in Kr.

One possible explanation is that the Fermi levels are not equalized. MgO and Kr are both insulators with large bandgaps of 7.8 and 12 eV, respectively. When there are no electrons available to align the Fermi levels, thermal equilibrium cannot be established and the vacuum levels will be aligned instead [161]. In addition, it is well known that the energy levels of rare-gas solids physisorbed on various metal surfaces are aligned to the vacuum level [162–164]. Therefore, we have also considered alignment of the vacuum levels. This is achieved analogous to the alignment procedure shown in Figures 38 and 39. Alignment of the vacuum levels is established by a potential shift $\Delta^{\mathrm{VAC}}_{\mathrm{align}} = \Delta_{\mathrm{MgO}} - \Delta_{\mathrm{Kr}}$ between the crystal zeros of the two materials. For positrons the shift is $-\Delta^{\mathrm{VAC}}_{\mathrm{align}}$, which automatically leads to the alignment of the vacuum levels for positrons, as expected. The numerical values are again listed in Table 11. The surface dipole and the positron work function of MgO, Δ_{MgO} and ϕ^{MgO}_+, respectively, were calculated with Eq. (50) using the experimental value for the electron work function of MgO, 3.8 eV [165]. The final result is that the difference between the positron ground states of the two materials is simply the difference between the positron work functions,

$$\Delta E^{\mathrm{VAC}}_{AB,+} = -(\phi^A_+ - \phi^B_+) \quad (56)$$

so that $-\phi_+$ takes the role of the positron affinity when the vacuum levels are aligned [compare with Eq. (53)]. With the values given in Table 11, this yields −0.5 eV for $\Delta E^{\mathrm{VAC}}_{AB,+}$, indicating that the nanocluster still acts as a (shallow) potential well. However, the differences in calculated energies

Table 11. Electronic and positronic data for MgO and Kr.

	MgO	Kr	Ref./Eq.
a_0 (Å)	4.212	5.72	[140, 142]
μ_- (eV)	−1.0	−13.3	[140, 142]
μ_+ (eV)	−5.0	−1.3	[140, 142]
ϕ_- (eV)	3.8	11.9	[142, 165]
ϕ_+ (eV)	2.2	2.7	Eq. (50), [142]
Δ (eV)	2.8	−1.4	Eq. (50)
E_c (eV)	6.8	−1.7	[140, 142]
E_g (eV)	7.8	11.6	[140, 142]
E_F (eV)[a]	2.9	−7.5	$\mu_- + 1/2E_g$
A^{METAL} (eV)	−6.0	−14.6	Eq. (52)
A^{INS} (eV)	−2.1	−8.8	Eq. (55)
$A^{\mathrm{VAC}} = -\phi_+$ (eV)	−2.2	−2.7	Eq. (56)

[a] Defined with respect to the crystal zero in Figure 38.

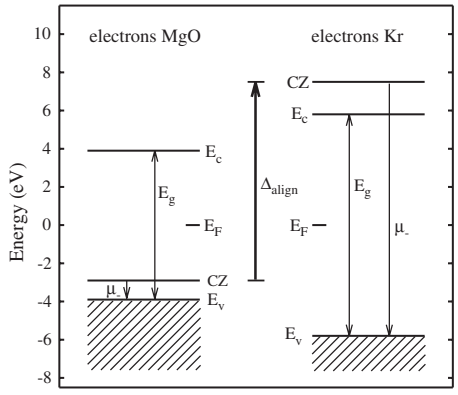


Figure 38. Electronic energy levels for MgO in contact with solid Kr, based on the data listed in Table 11. The alignment of the Fermi levels is achieved by an energy shift $\Delta^{\mathrm{INS}}_{\mathrm{align}} = (\mu^A_- + \frac{1}{2}E^A_g - \mu^B_- - \frac{1}{2}E^A_g)$ between the crystal zeros of the two materials.

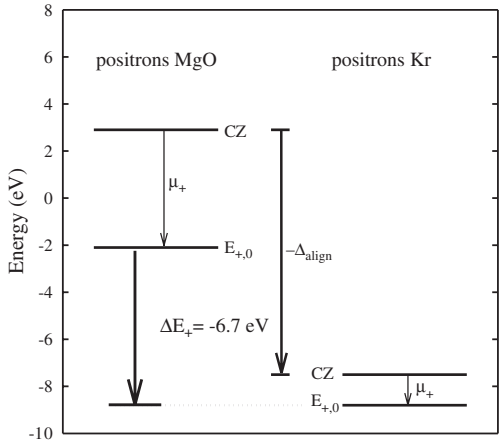


Figure 39. Positronic energy levels for MgO in contact with solid Kr, based on the data listed in Table 11. $\Delta^{\mathrm{INS}}_{\mathrm{align}}$ is defined in Figure 38.

between various theoretical models are approximately 1 eV. Furthermore, we have combined experimental and theoretical results, and the electronic and positronic data for Kr given in Table 11 are only valid for a lattice parameter of 5.72 Å while in reality, there is a distribution. Therefore, we estimate the systematic error in $\Delta E_{AB,+}$ to be at least 1.5 eV, so that it is not clear whether positrons will trap in Kr nanoclusters when the vacuum levels are aligned. The various possibilities for alignment of the electronic energy levels and the corresponding differences in positronic ground potentials are summarized in Table 12.

One more aspect to be considered is that Kr is commonly used as a moderator because positrons thermalize ineffectively in solid rare gases [166]. This means that positrons hardly lose energy when moving in "bulk" solid Kr. So even if the Kr nanoclusters would act as a potential well, a positron entering a nanocluster might not lose enough energy to be thermalized and to remain trapped inside the cluster until annihilation. In the literature, trapping in spherical Kr bubbles in metal (Ni, Cu) is reported [143, 144] and these authors conclude that the trapping occurs at the interface rather than in the nanocluster itself. We cannot fully exclude that positrons are also trapped at the MgO//Kr interface, but it is not likely. First, the TEM results suggest that the cubically shaped Kr clusters fit very well in the MgO lattice, reducing the probability of interface defects that can act as trapping sites. Second, if the positrons would trap at the interface, one would still expect an anisotropic contribution from Kr to the 2D-ACAR spectrum due to overlap of the positron wave functions with the solid Kr. Unfortunately, the anisotropy in the electronic/positronic structure for bulk Kr is not known theoretically nor experimentally, so that we cannot interpret our experimental results in more detail.

Summarizing, the experimental 2D-ACAR spectrum does not show any anisotropic contribution from Kr, showing that positrons do not trap in Kr nanoclusters. Three effects can be responsible for this. It is likely that the Fermi levels will not equalize due to the excellent insulating properties of both materials. Therefore, it is possible that the nanoclusters do not act as a positronic potential well (Table 12). Positrons do not thermalize in solid Kr, so probably the positrons would leave the cluster even if it would act as a potential well. Trapping at the interface is not likely as the cubically shaped Kr clusters fit very well in the MgO lattice. Moreover, if the positrons would trap at the interface, an anisotropic contribution from Kr to the 2D-ACAR spectrum would be expected due to overlap of the positron wave function with the solid Kr.

Table 12. Positronic potential difference between Kr cluster and MgO in contact assuming alignment of the Fermi levels in metals [Eq. (51)], in insulators [Eq. (54)], or assuming alignment of the vacuum levels [Eq. (56)].

Level alignment	$\Delta E_+ = E_+^{\mathrm{Kr}} - E_+^{\mathrm{MgO}}$	Result (eV)
Fermi (metals)	$\Delta E_+^{\mathrm{METAL}} = (\mu_+^{\mathrm{Kr}} + \mu_-^{\mathrm{Kr}}) - (\mu_+^{\mathrm{MgO}} + \mu_-^{\mathrm{MgO}})$	N.A.[a]
Fermi (insulators)	$\Delta E_+^{\mathrm{INS}} = (\mu_+^{\mathrm{Kr}} + \mu_-^{\mathrm{Kr}} + \frac{1}{2}E_g^{\mathrm{Kr}}) - (\mu_+^{\mathrm{MgO}} + \mu_-^{\mathrm{MgO}} + \frac{1}{2}E_g^{\mathrm{MgO}})$	-6.7 ± 1.5
Vacuum	$\Delta E_+^{\mathrm{VAC}} = (-\phi_+^{\mathrm{Kr}}) - (-\phi_+^{\mathrm{MgO}})$	-0.5 ± 1.5

[a] Does not apply to the MgO//Kr insulator–insulator interface.
Note: The data of Table 11 have been used to calculate the numerical values.

5.4. Discussion

Solid Kr nanoclusters have been successfully created in MgO by means of 280 keV Kr ion implantation and subsequent thermal annealing at 1100 K. The nanoclusters have been observed by X-TEM in high-resolution mode. The fcc Kr nanoclusters are rectangularly shaped with sizes of 2 to 4 nm. The lattice parameter of the solid Kr in the clusters varies from 5.3 Å for small clusters to 5.8 Å for large clusters, with the lattice parameter increasing with increasing nanocluster size. Using the Ronchi EOS, this corresponds to local pressures of 0.6–2.5 GPa. Both optical absorption spectroscopy and Doppler broadening positron beam analysis show that small defects are present in the MgO after the ion implantation and that defect aggregation occurs during annealing. Experimentally, no evidence was found for positron confinement in Kr nanoclusters. The nonequalization of Fermi levels, the poor thermalization of positrons in solid Kr, and the absence of trapping sites at the Kr//MgO interface might be the reason that positrons are not trapped in Kr clusters.

6. NANOVOIDS IN He IMPLANTED AND THERMALLY ANNEALED MgO

6.1. Introduction

Interest in the behavior of helium in MgO has two diverse origins. The first is analogous to the case discussed in the previous section, namely that MgO acts as a host material for the generation of nonlinear and other optical properties by introducing nanometer-sized metallic or semiconductor clusters. Control over the size and shape of these nanoprecipitates might be obtained by a two-step process when metallic ions are introduced in previously generated nanometer-sized voids in MgO. It is suggested that the cavities will act as gettering centers to metallic ions as has been observed for nanocavities in silicon [167, 168]. The second application for MgO is as a candidate uranium-free host matrix for the transmutation of actinides generated in nuclear reactor fuels. Alpha decay of actinides introduces helium atoms into the MgO matrix that interact with radiation damage defects. It is therefore of interest to investigate the relevant activation energies of helium dissociation from various defects where the helium might be trapped. In this section we will discuss the shape of the cavities as examined with TEM.

6.2. Experiments

Samples obtained by cleaving of MgO(100) crystals to a thickness of 1 mm and a size of 1×1 cm^2 were implanted with a dose varying from 1×10^{19} to 1×10^{20} m^{-2} ^{3}He ions with 30 keV energy. Thermal annealing of the samples was performed in a longitudinal oven in ambient air during 30 minutes. After each annealing step, the samples were analyzed by photon absorption and positron beam analysis.

Photon absorption revealed the presence of F-centers in the as-implanted sample, which disappeared after annealing to 500 K. For the high dose implanted samples annealed at 1370 K positron beam analysis indicated the presence of voids with nanometer size at a depth somewhat shallower than of the helium implantation depth [169]. In a similar experiment it was shown using neutron depth profiling that helium is released from the cavities at a temperature of about 1300 K.

For cross-sectional TEM examination slices with a thickness of about 1 mm were obtained from the MgO single crystal with the voids using cleavage fracture along {100}. Two slices were glued together with the original surfaces exposed to implantation facing each other and separated by the glue (Gatan G-1 epoxy and hardener). The coupled slices of MgO were glued in cross-section geometry on a Cu ring (3 mm outer and 2 mm inner diameter). Then the MgO was ground and polished to a thickness of about 100 μm. Using dimpling from the Cu side the thickness of the MgO in the center of the ring is reduced to about 25 μm. Finally, the dimpled sample was thinned to electron transparency by two beams of 4 kV Ar^+ ions making an angle of 6° with the top and bottom sample surfaces using a Gatan PIPS 691. TEM observations were performed using a JEOL 4000 EX/II operating at 400 kV (point to point resolution of 0.165 nm).

An overview of the implanted top layer of the MgO is shown in the bright-field TEM image of Figure 40 (viewing direction close to [001] and in a two-beam condition with the other strong beam $g = 220$). At a distance of about 50–60 nm from the edge of the foil a localized band of rectangular features can be observed. These rectangular features are visible due to Fresnel contrast resulting in a bright outer fringe and a dark inner fringe indicating the overfocus condition during imaging. Figure 41 shows the rectangular features in underfocus condition. The Fresnel contrast makes clear that rectangularly shaped voids, due to the He implantation and the annealing treatments, are present in the MgO

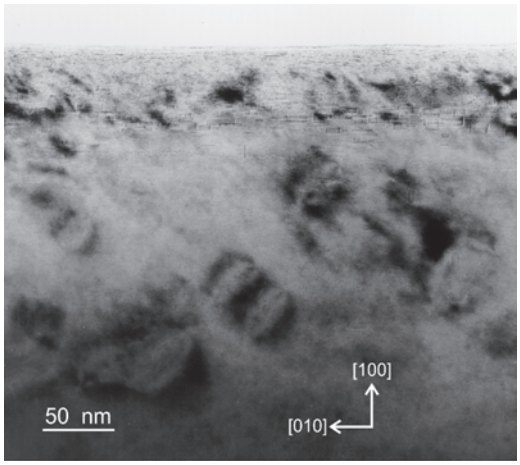


Figure 40. Bright-field TEM image (two-beam condition with other strong beam $g = 220$) showing an overview of the implanted top layer revealing the presence of sharply rectangular nanocavities as seen by Fresnel contrast in overfocused condition.

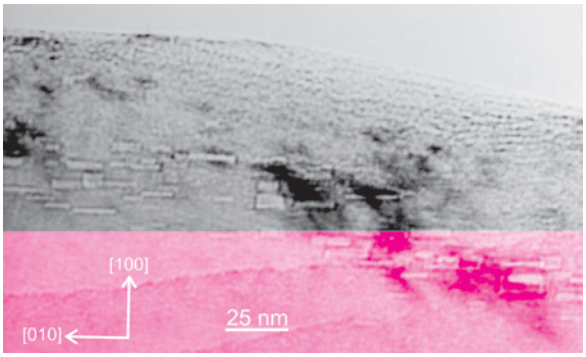


Figure 41. Bright-field TEM image (two-beam condition with viewing direction near to [001] and with other strong beam $g = 220$) showing sharply rectangular nanocavities as seen by Fresnel contrast in underfocused condition.

with observable sizes of the edges ranging between 0 and 20 nm. The crystal orientation of the MgO is well defined (with the edge of the TEM specimen still more or less parallel to the original implanted surface) and the sides of the rectangular voids correspond to cube planes of the MgO. The present results clearly indicate that already for voids in the nanometer scale range in MgO a very strong tendency for faceting is present. This is uncommon for nanoscale voids (bubbles) obtained after He implantation and annealing in metals or semiconductors and previously investigated ceramics such as α-Al_2O_3 [170–174]. Then, the nanovoids are spherical and only for increased sizes (e.g., larger than a few tens of nm) do they become faceted.

The origin of the strong tendency for faceting in the MgO system is most probably related to the high (Pauling's) ionicity of MgO. The surface energy for different facets of MgO is largely determined by the polarity of the terminating plane. Nonpolar surfaces have finite energies whereas in case of a dipole moment in the repeat unit perpendicular to the surface the energy diverges and becomes infinite [175]. Hence, for highly ionic systems spherical voids are even at a size of a nanometer unfavorable, because locally an uncompensated dipole moment perpendicular to the surface may result.

The shape of the voids is expected to be independent of its size when the Wulff construction is applied. However, for decreasing size of cavities the fraction of atoms on or near edges and also of atoms in or close at corners that contribute to the total energy of the cavity increases. Hence, the prediction of the shape by surface energy of individual facets alone will no longer apply. Minimization of the surface area with respect to the enclosed volume tends to have a controlling influence on the shape of the void. For metals and semiconductors and ceramics aforementioned this mechanism, which leads to small spherical cavities, applies, but for voids in MgO it does not. Apparently, due to the extreme anisotropy in energy of atoms in different surface states, the voids in MgO are bounded by cube planes of the MgO, independent of the size of the cavities. In the literature, surface energies for MgO (111) and (110) are quoted to be a factor 2 to 4 times larger than for the (100) surface [176].

In the region within 50–60 nm from the edge in Figure 41 hardly any cavities are found indicating the locally low He concentration after implantation. Helium in this region can have formed undetectably small clusters or can have diffused to earlier nucleated cavities at larger depth or to the surface followed by desorption. At a larger distance from the edge than where the cavities are located, other sharply contrasting features can be observed in Figure 42. These features become more apparent at a distance of about 250 nm from the edge in Figure 42, showing a strong-beam dark-field image for viewing near to [001] for $g = 420$ as indicated by the arrow. Note that the dark-field image is inverted for reasons of clarity (i.e., the gray scale from black to white is white to black instead). These features can be identified as dislocation loops, often with the same inclination to the viewing direction. Dislocation loop contrast may also originate from (partially) coherent precipitates [177]. However, in the present context of He implantation and annealing the presence of dislocation loops is not surprising whereas of precipitates it is unlikely. Dislocation loops in MgO as a result of irradiation damage followed by annealing were found to be of interstitial type with a Burgers vector of $1/2\langle 110\rangle$. The low density and large sizes of the dislocation loops compared to the cavities indicate the low concentration of implanted He and large diffusion distances at the depth of the dislocation loops.

The white dashed line in Figure 42 indicates the center of the region where the cavities are localized; although not visible in Figure 42 they could be well recognized on the original negative. For determination of the depth of the cavities and the dislocation loops below the original sample surface Figure 40 is not suitable, because during preparation of the TEM sample a part of the top layer has been removed by ion milling. On the other hand, Figure 42 is suitable for obtaining the depth of the voids and dislocation loops since the region imaged in Figure 42 is very close to unremoved glue between the surfaces of the two MgO slices. Therefore, we can be sure that the cavities are located at a depth of about 120 nm (almost all localized in between 100 and 140 nm) and the dislocation loops are localized at a depth of 150–350 nm (mostly at a depth of about 250 nm).

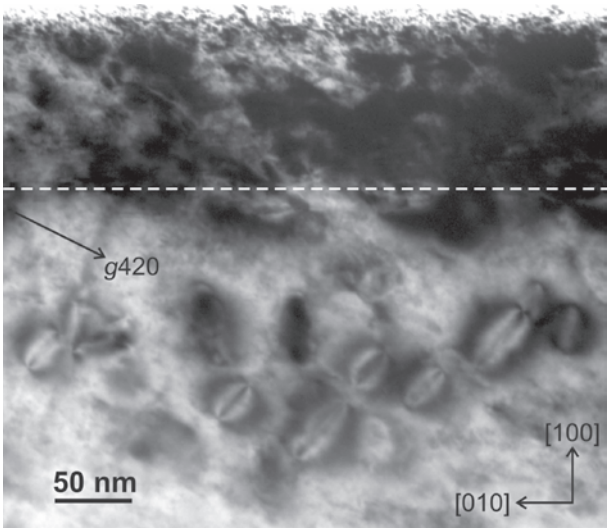


Figure 42. Strong-beam, dark-field TEM image for a viewing direction near to [001] with $g = 420$ showing dislocation loop contrast at a depth between 150 and 300 nm. The white dashed line at a depth of 120 nm indicates the center of the region where the rectangular voids are present (they are not visible in the present micrograph but clearly traceable in the original negative). For reasons of clarity the intensity scale of the image is inverted (i.e, the gray scale from black to white is white to black instead).

As mentioned earlier positron beam analysis had been performed to the sample before TEM examination. Positrons have a high affinity for trapping at vacancies and agglomerates of vacancies or voids. In general the 511 keV photon peak observed due to annihilation of the positron with an electron from the sample is Doppler broadened. The electron carries a certain momentum component to or from the detector so that broadening occurs. The momentum of the electrons encountered in the defects is generally lower than in the bulk and therefore broadening is reduced in that case. When the cavity is larger than about 0.8 nm positronium, the bound e^+–e^- pair is formed in the cavity which annihilates with even less broadening. The shape of the peak is characterized quantitatively with the aid of two parameters S and W. S accounts for the weight of the low momentum contributions and W for the high momentum contributions (see Section 2). In this section we will use the S-parameter versus positron implantation energy measurements to investigate the depth profile of the formed cavities.

The positron results are shown in Figure 43, where the S versus E curves obtained for several annealing temperatures are plotted. By fitting the results with a three-layer model

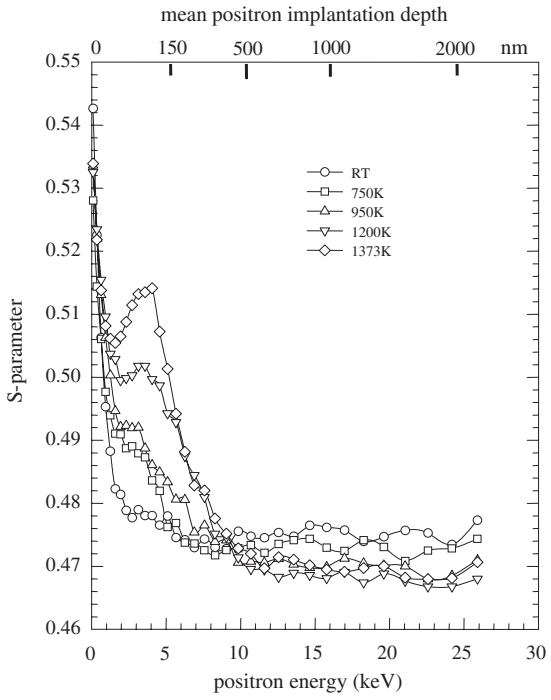


Figure 43. The S-parameter as a function of positron energy for the 1×10^{16} cm^{-2} ^3He implanted MgO(100) crystal after annealing at the indicated temperatures. The mean probe depth of the positrons is indicated in the top scale. Note the increase of the S-parameter upon annealing at the depth corresponding with the cavity layer from Figure 42.

allowing for diffusion and defect trapping of the positrons two defect layers are identified on the substrate. The first layer is the layer where according to TRIM calculations displacement damage without helium is produced by the helium implantation. During annealing the *S*-parameter increases gradually from a value of 0.478 to a saturation value of 0.505 at 1200 K. This indicates that some vacancy clustering occurred in the layer. In the second layer, where the helium was implanted and displacement damage reached its maximum, the bubbles were formed which after helium release turned into the observed cavities. The fitted *S*-parameter of 0.523 found for that layer clearly indicates the formation of positronium in line with the microscopic observation. The zone of depth where the cavities are found coincides rather well with the defect layer 2. The best fit was obtained for a zone from 100 to 140 nm. For the sample with half the helium implantation dose of the TEM sample the annealing was continued to higher temperature. It was observed that the cavities disappeared at 1500 K. At that temperature self-diffusion is a sufficiently fast process to let shrinkage of the cavities occur. We explain the formation of the cavities by the stabilization of helium during the early stages of vacancy clustering. Helium prevents the vacancy clusters from collapsing to dislocation loops; therefore three-dimensional clusters are formed. The growth of the helium filled vacancy clusters proceeds by an Ostwald ripening process: small clusters will dissociate and larger clusters or bubbles will grow by trapping the helium and vacancies from the dissociated clusters. The final stage is the release of helium from the bubbles by permeation through the top layer of the material.

It appears that a temperature window exists from 1200 to 1400 K in which helium is permeating while the cavities remain thermally stable. In earlier work [172] where MgO was neutron irradiated, cavities or voids were never observed. Apparently, it is essential to have a sufficiently high gas concentration in MgO to promote the void formation. It is well known that for a variety of ceramic materials (e.g., Al_2O_3 exposed to neutron irradiation) ordered void structures are developed. It is likely that gases, notably transmutation produced helium, are involved which may stabilize the growing voids. In contrast to metals, where helium remains bound and stabilizes the cavities or bubbles to temperatures close to melting, in ceramics helium permeates from the cavities to the vacuum at a much lower temperature and leaves cavities behind which are stable to temperatures where they dissolve by vacancy dissociation. If the implanted atom has a higher solubility than helium or hydrogen, nanoprecipitates in the case of metal atoms or gas bubbles in the case of the heavier noble gas atoms are observed. Gold precipitates have been observed in MgO [178] and gas bubbles of nanosize have been observed in UO_2 [179] and in MgO [171]. Larger gas bubbles of micrometer size are known from MgO crystals grown from the melt [180].

This section has shown that rectangular and flat nanosize cavities are formed, which have a thickness generally much smaller than the lateral dimensions. In the future it is of interest to know what morphology the cavities will have when lower helium doses are used. It has already been shown that existing cavities can be grown larger by postirradiation with hydrogen followed by thermal annealing. Of importance for the study of interfaces will be the possibilities offered to study interfaces and interfacial reactions by observing the cavities and their shape by refilling the cavities with gas or other atoms. The cavities form an interesting materials-testing laboratory in a nanoworld protected from the environment.

6.3. Calculations of the Vacancy Cluster Structures in MgO and Their Interaction with Helium

6.3.1. Methodology

It is important to study the nucleation of voids by calculating the formation energies of small vacancy clusters. Defect formation energies are compared with calculations and experimental data from literature [156, 181, 182]. Furthermore, the activation energies associated with helium migration and dissociation from vacancy-type defects are calculated. These energies are compared with the activation energy for permeation of helium through a 100 nm thick layer of (100) MgO as observed with NDP.

Calculations have been carried out using the Mott–Littleton methodology as implemented in the code CASCADE [183]. Atomistic simulations are based on a classical Born description of the lattice with pair potentials acting between atomic species. The ionic interactions have two components: a long-range Coulombic interaction and a short-range interaction, which accounts for the effect of electron cloud overlap and van der Waals energies. In this study short-range terms are approximated using parameterized pair potentials of the Buckingham form where A, ρ, and C are the potential parameters:

$$S(r_{ij}) = A_{ij}e^{-r_{ij}/\rho} - C_{ij}/r_{ij}^6 \tag{57}$$

Two distinct potential parameter sets are used in this study, the parameters of which are listed in Table 13. The potential set for the so-called full charge model was derived with the assumption that the lattice ions assume their full charge states of $+2.0e$ and $-2.0e$ for Mg and O respectively. The potential set for the partial charge model was derived for Mg and O ions with a partial charge of $+1.7e$ and $-1.7e$ [184]. This partial charge state is based on quantum mechanical calculations. The parameters of both potential sets were derived by simultaneous fitting to the relevant crystal properties (e.g., lattice parameter, elastic and dielectric constants) [184]. In addition to the pair potentials, the Dick and Overhauser shell model is used to describe polarization effects. The shell model parameters are also listed in Table 13. The method is further explained in [185]. The variations of helium–magnesium and helium–helium energies as a function of atomic separation were calculated within the electron-gas approximation [186] with subsequent fit to a Born–Mayer potential form. The important helium–oxygen potential was taken from [187], where energies were calculated within the Hartree–Fock approximation based on a distribution of oxygen ions characteristic of the UO_2 lattice.

Throughout this section, defects will be written in a modified Kröger–Vink notation (see Table 15). Activation energies for migration of vacancies and interstitials are obtained

Table 13. Potential parameters for the full and partial charge model.

			Full charge model	Partial charge model	Ref.
Short-range potentials					
O–O	Buckingham	A (eV)	9547.96	2,230,386.3	[184]
		ρ (Å^{-1})	0.21916	0.1429	
		C (eV Å^6)	32	32	
Mg–O	Buckingham	A	1284.38	959.3	[184]
		ρ	0.29969	0.2994	
		C	0	0	
He–He	Buckingham	A	166.8	166.8	*
		ρ	0.28096	0.28096	
		C	0	0	
He–Mg	Buckingham	A	975.5	975.5	*
		ρ	0.2229	0.2229	
		C	0	0	
He–O	Lennard-Jones	A	2247.836	2247.836	[185]
		B	11.762	11.762	
Shell model parameters					
O shell charge			−2.04	−2.23	[184]
O core charge			0.04	0.53	
O total charge			−2	−1.7	
Mg total charge			2	1.7	
O harmonic constant		K (eV Å^{-2})	6.3	32.2	

* The He–He and Mg–He energies were calculated within the electron-gas approximation and subsequently fitted to a Born–Mayer potential form.

by calculating defect formation energies at discrete positions along the migration path and subtracting the minimum formation energy at the interstitial position from the maximum formation energy at the saddle point (see Fig. 44). The migration paths for charged magnesium and oxygen vacancies V''_{Mg}, $V_O^{\bullet\bullet}$, and charged magnesium and oxygen interstitials $Mg_i^{\bullet\bullet}$, O_i'' follow straightforwardly from the simple structure of MgO [181]. Interstitial helium atoms $He_i^{\times}$ follow the same migration path as charged oxygen and magnesium interstitials, in a ⟨100⟩ direction (see also Fig. 47).

6.3.2. Results and Discussion

First, the formation energies of Schottky and Frenkel defects, the migration activation energies of oxygen and magnesium vacancies and interstitials, as well as the migration activation energy of helium interstitials in MgO were calculated. The results are listed in Table 14. The energies reported here show good agreement with earlier INDO and pair potential calculations [181, 182] and with experimental results. In general, values reported in the literature are in between the full charge and partial charge values of Table 14. Although the absolute values calculated with the partial charge model agree better with the experimental data, the relative energies predicted by each model are essentially the same. Thus, for example, Schottky disorder is proportionally more favored over Frenkel disorder by both models.

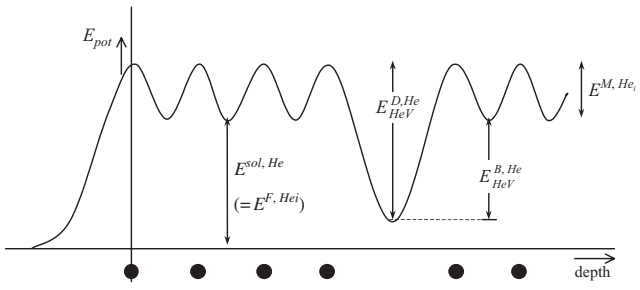


Figure 44. Potential energy diagram of helium interactions with a solid. The peaks in the potential energy diagram correspond to the location of a plane of atoms (indicated by the atoms in the bottom of the figure).

Table 14. Formation energies of Schottky and Frenkel defects, migration energies of oxygen and magnesium vacancies and interstitials, and the formation and migration energy of helium in MgO.

Defect energies (eV)		Full charge model	Partial charge model
Lattice energy	E^{latt}_{MgO}	−41.16	−29.74
Schottky pair	$E^{F,\,V_O^{\bullet\bullet}} + E^{F,\,V''_{Mg}}$	8.44	6.12
Schottky cluster	$E^{F,\,\{V_{Mg}V_O\}^{\times}}$	5.35	3.90
Frenkel–Mg	$E^{F,\,V''_{Mg}} + E^{F,\,Mg_i^{\bullet\bullet}}$	13.35	9.58
Frenkel–O	$E^{F,\,V_O^{\bullet\bullet}} + E^{F,\,O_i''}$	13.81	9.95
Migration energies	$E^{M,\,V''_{Mg}}$	2.08	1.53
	$E^{M,\,V_O^{\bullet\bullet}}$	2.12	1.56
	$E^{M,\,Mg_i^{\bullet\bullet}}$	0.64	0.46
	$E^{M,\,O_i''}$	0.58	0.47
	$E^{M,\,He_i}$	1.08	0.71
He solution energy	$E^{\text{sol,He}} = E^{F,\,He_i}$	3.19	2.80

Note: Definitions of the formation and migration energies can be found in Figure 44.

Vacancy Clusters Next, the energetically most favorable configurations of selected vacancy clusters composed of up to eight vacancies were determined. A number of possible configurations are displayed in Figure 45 and the corresponding formation energies are listed in Table 15. Again both partial and full charge models yield equivalent relative geometric preferences.

By considering Figure 45 and Table 15 it is clear that configurations which induce low polarization around the vacancy cluster are energetically favorable. An interesting extension of this hypothesis is that the energetically most favorable configuration for an eight-vacancy cluster is a cube. Here a three-dimensional configuration is preferred to any planar configuration. A vacancy cluster consisting of a monolayer "slab" induces polarization oriented perpendicular to the monolayer slab. However, when there are two adjacent monolayers of vacancies, the vertically orientated polarization induced by the two monolayers cancels to a large degree due to the alternate arrangement of magnesium and oxygen atoms in the MgO matrix. Therefore three-dimensional growth in an even number of slabs reduces the formation energy of vacancy clusters. This explains why voids in MgO are relatively easy to create.

Helium Defect Interactions Now that the configurations of the energetically most favorable vacancy clusters have been determined, an approximation for the activation energies for dissociation of the nth helium atom from a He_nV_m vacancy-type defect (i.e., $V_mHe_n \rightarrow V_mHe_{n-1} + He_i$) can be calculated. Here again V_m can be either an oxygen or magnesium monovacancy or a cluster composed of an equal number of oxygen and magnesium vacancies, totaling m vacancies. The dissociation energies $E^{D,\mathrm{He}}_{\mathrm{He}_n\mathrm{V}_m}$ have been calculated by summing the binding energy of the helium atom in the vacancy cluster, $E^{B,\mathrm{He}}_{\mathrm{He}_n\mathrm{V}_m}$, and the migration energy of interstitial helium in defect-free MgO, $E^{M,\mathrm{He}}$. The binding energy is calculated from formation energies of helium-filled vacancy clusters, $E^{F,\mathrm{He}_n\mathrm{V}_m}$, and the helium interstitial formation energy, E^{F,He_i}.

$$E^{D,\mathrm{He}}_{\mathrm{He}_n\mathrm{V}_m} = E^{B,\mathrm{He}}_{\mathrm{He}_n\mathrm{V}_m} + E^{M,\mathrm{He}} \tag{58}$$

$$E^{B,\mathrm{He}}_{\mathrm{He}_n\mathrm{V}_m} = E^{F,\mathrm{He}_{n-1}\mathrm{V}_m} + E^{F,\mathrm{He}_i} - E^{F,\mathrm{He}_n\mathrm{V}_m} \tag{59}$$

The definitions of these energies can be found in Figure 44. The calculated dissociation activation energies are listed in Table 16 and are depicted in Figure 46.

It is clear from Table 16 and Figure 46 that the dissociation energy of helium from a vacancy or vacancy cluster containing only one helium atom ($n = 1$) is more or less constant: 4.4 eV for the full charge model and 3.6 eV for the partial charge model. Helium atoms experience empty vacancy clusters as vacuum and therefore the dissociation energy does not increase with larger cavities. Only the magnesium vacancy is apparently somewhat smaller than the oxygen vacancy. Another result of importance to note is that the fifth helium atom in an oxygen vacancy ($n = 5$) is not bound at all; that is, the dissociation energy for the fifth helium atom in a $\{V_OHe_5\}^{\bullet\bullet}$ defect is less than the helium migration energy in the MgO bulk (Tab. 14, Fig. 46) for both the full charge and partial charge model. Thus the binding energy of the cluster is negative and such a cluster would not be stable.

The activation energy for helium atom dissociation from a cluster as we have calculated assumes that the migration activation energy for a helium atom, $E^{M,\mathrm{He}}$, as it moves away from the cluster, is the same as in the bulk defect-free material. However, local relaxation of the lattice surrounding the defect may increase (or decrease) this energy.

In order to learn more about the dissociation of a single helium atom from a monovacancy, the potential energy of the helium atom has been calculated along the escape trajectory from an oxygen vacancy as depicted in Figure 47. From here on, only the partial charge model is considered. The helium potential along the trajectory as a function of the distance from the center of the monovacancy is shown in Figure 48. Regarding the saddle points that can be distinguished in Figure 48, it is clear that the potential energy at the first saddle point is slightly larger than at the subsequent saddle points. It results in a dissociation energy of 3.9 eV rather than 3.5 eV [obtained with Eqs. (58) and (59)]; see Table 16. This effect is the result of the charged nature of the monovacancies: an oxygen monovacancy exhibits a charge of $+1.7\,e$ with respect to the defect-free MgO matrix. The nearest-neighbor magnesium atoms will be repelled by Coulombic interactions and the second-neighbor oxygen atoms will be attracted to the defect. The expansion and contraction of the first and second ion shells around the oxygen vacancy result in a potential barrier, which increases the activation energy of helium dissociation from a monovacancy. This effect (although without helium) has already been reported by Kotomin and Popov [182]. The increased activation energy for the monovacancy is expected to be

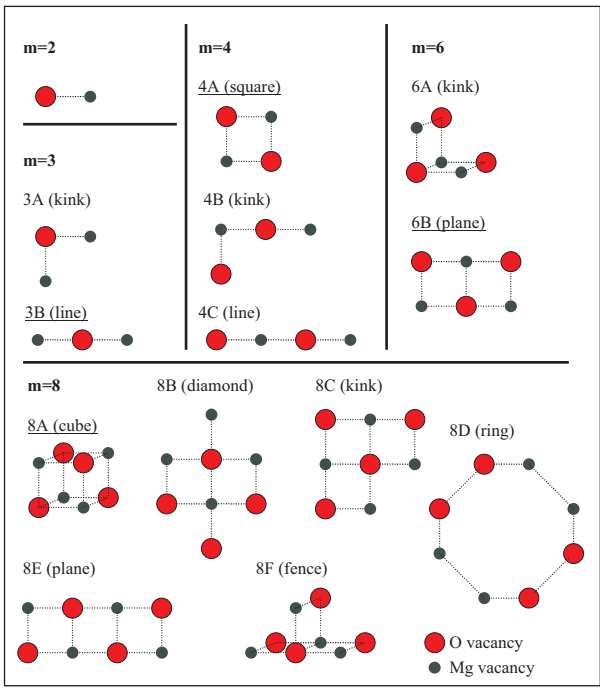


Figure 45. Possible vacancy cluster (V_m) configurations. Note that the depicted atoms represent vacancy sites. The formation energies are listed in Table 15. The energetically most favorable configurations are underlined.

Table 15. Formation energies of vacancy clusters V_m, $1 \le m \le 8$ (here m is the total number of vacancies in the cluster: either oxygen, magnesium, or both).

Number of vacancies m	Configuration of cluster V_m (see Fig. 45)		Formation energies (eV)	
			Full charge model	Partial charge model
Charged clusters				
$m = 1$	V''_{Mg}		3.93	2.63
	$V_O^{\bullet\bullet}$		4.51	3.49
$m = 3$	$[\{V_{Mg}\}_2\ \{V_O\}_1]''$	3A: kink	7.85	5.51
	$[\{V_{Mg}\}_2\ \{V_O\}_1]''$	3B: line	7.31	5.10
	$[\{V_{Mg}\}_1\ \{V_O\}_2]^{\bullet\bullet}$	3A: kink	8.42	6.37
	$[\{V_{Mg}\}_1\ \{V_O\}_2]^{\bullet\bullet}$	3B: line	7.88	5.97
Uncharged clusters				
$m = 2$	$\{V_{Mg}V_O\}^{\times}$	Schottky cluster	5.35	3.90
$m = 4$	$\{V_{Mg}V_O\}_2^{\times}$	4A: square	7.48	5.48
	$\{V_{Mg}V_O\}_2^{\times}$	4B: kink	9.19	6.72
	$\{V_{Mg}V_O\}_2^{\times}$	4C: line	9.12	6.66
$m = 6$	$\{V_{Mg}V_O\}_3^{\times}$	6A: kink	10.69	7.84
	$\{V_{Mg}V_O\}_3^{\times}$	6B: plane	9.96	7.30
$m = 8$	$\{V_{Mg}V_O\}_4^{\times}$	8A: cube	10.22	7.52
	$\{V_{Mg}V_O\}_4^{\times}$	8B: planar diamond	16.46	12.08
	$\{V_{Mg}V_O\}_4^{\times}$	8C: planar kink	12.91	9.49
	$\{V_{Mg}V_O\}_4^{\times}$	8D: ring	23.56	17.17
	$\{V_{Mg}V_O\}_4^{\times}$	8E: rectangle	12.36	9.07
	$\{V_{Mg}V_O\}_4^{\times}$	8F: fence	14.28	10.47

Note: The energetically most favorable configurations are underlined.

higher than for any of the neutral clusters since distortions to the surrounding lattice by neutral defects are much less.

Helium Permeation through MgO The calculated values of the helium dissociation energy can be compared with the results of helium permeation from cavities in MgO. A dense layer of helium bubbles at a depth of 100 nm was created by ^{3}He ion implantation in a single crystal of (100) MgO and subsequent annealing at a temperature of 1200 K [188]. By NDP it is observed that helium permeates to the surface during isochronal annealing at a temperature of 1323 ± 70 K on a time scale of 30 min. Cross-sectional transmission electron microscopy investigations have shown that the 100 nm thick layer between the cavities area and the surface is reasonably defect-free [120].

When dissolving the helium is the rate-determining step, the helium release from the layer of gas bubbles can be modeled as a quasi-stationary flow of gas. This flow is sustained by the gradient that exists from the equilibrium concentration $c_{He,\,eq}$ of dissolved gas near the cavity layer to the

Table 16. Dissociation energies for the nth helium atom in $V_{Mg}He_n$, V_OHe_n, and V_mHe_n defects for the full charge and partial charge model.

	Full charge model—dissociation energies (eV)					
n	V''_{Mg}	$V_O^{\bullet\bullet}$	V_2 (Schottky)	V_4 (square)	V_6 (plane)	V_8 (cube)
1	4.21	4.30	4.37	4.39	4.38	4.37
2	2.01	2.99	4.08	4.26	4.37	4.37
3	1.98	2.33	3.06	4.12	4.22	4.26
4	2.49	2.50	2.61	3.99	4.22	4.27
5	1.60	0.72	2.83	3.09	4.02	4.09
6	2.09	1.52	2.35	0.87	3.82	4.05
	Partial charge model—dissociation energies (eV)					
n	V''_{Mg}	$V_O^{\bullet\bullet}$	V_2 (Schottky)	V_4 (square)	V_6 (plane)	V_8 (cube)
1	3.46	3.54	3.61	3.63	3.62	3.61
2	1.47	2.27	3.32	3.50	3.61	3.60
3	1.46	1.69	2.16	3.36	3.46	3.50
4	1.97	1.87	2.12	3.23	3.46	3.38
5	1.14	0.37	2.16	2.37	3.26	3.46
6	1.56	1.06	1.06	2.21	3.07	3.30

Note: Energies have been calculated using Eqs. (58) and (59) using the energetically most favorable configuration for the vacancy clusters listed in Table 15. The data are displayed in Figure 46.

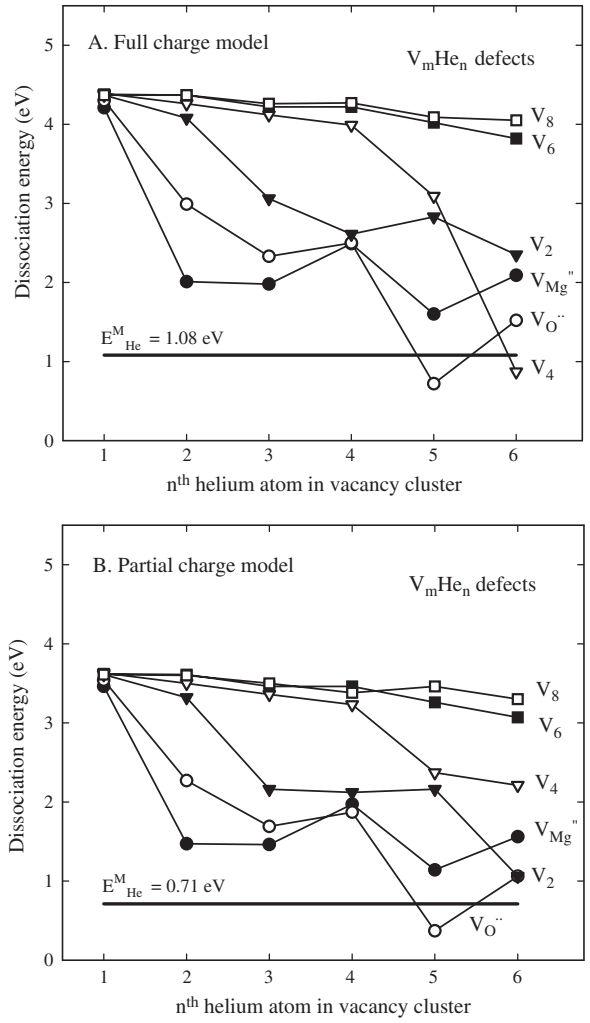


Figure 46. Activation energies for the dissociation of one helium atom from helium-filled vacancy clusters (i.e., $V_mHe_n \rightarrow V_mHe_{n-1}$ + He_i) as a function of the number of helium atoms n in the cluster prior to dissociation. Lines are drawn to guide the eye.

zero concentration near the external surface,

$$\frac{dN_{He}}{dt} = -D\frac{dc_{He}}{dx} = -D\frac{c_{He,\,eq}}{L} \tag{60}$$

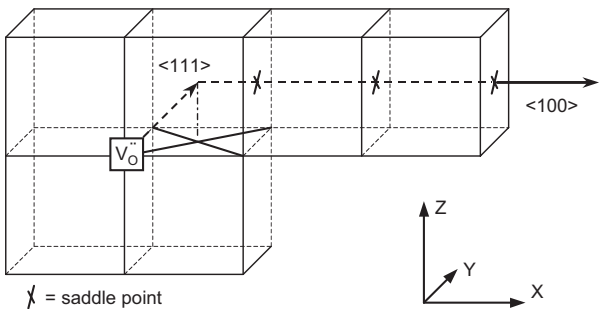


Figure 47. Assumed trajectory of a helium ion dissociating from an oxygen monovacancy.

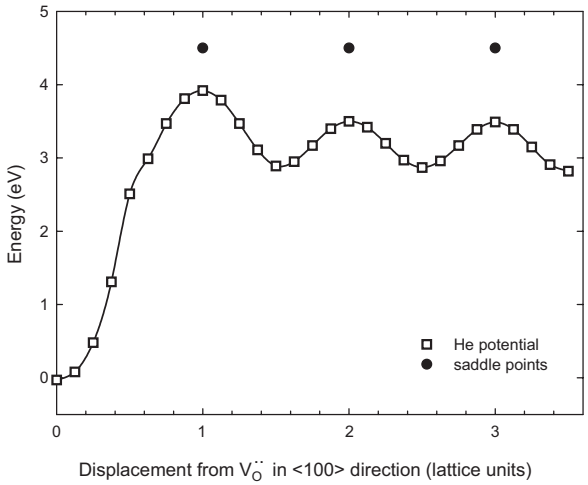


Figure 48. Potential energy graph of a helium atom dissociating from an oxygen vacancy in MgO, following the trajectory of Figure 47. Only the partial charge model is considered.

where dN_{He}/dt is the flow rate of helium atoms per unit area in m^{-2} s^{-1}, c_{He} is the concentration of helium atoms in m^{-3}, x is the depth in the sample in m, and L the depth of the cavity layer in m. D is the helium diffusivity in m^2 s^{-1} and is related to the helium migration energy E_M as [189]

$$D = \frac{1}{6}\lambda^2\frac{\omega}{2\pi}\exp(-E_M/kT) \tag{61}$$

Here λ is the jumping distance between two interstitial sites (2.1 Å), ω is a fundamental attempt frequency taken as 10^{13} s^{-1}, k is the Boltzmann constant (8.62×10^{-5} eV K^{-1}), and T is the temperature in K. Substituting 0.71 eV for E_M (Table 14) in Eq. (61), the relaxation time to establish a stationary flow $\tau_D = L^2/2D$ is of the order of 10^{-4} s, orders of magnitude smaller than the gas release time (of the order of minutes). The approach of quasi-stationary flow is therefore feasible. The equilibrium concentration of helium in MgO is [189]

$$c_{He,\,eq} = c_{MgO}\ \exp(\mu/kT - G_{sol}/kT) \tag{62}$$

with c_{MgO} the atomic concentration of MgO, 1.07×10^{29} atoms m^{-3}. μ is the chemical potential and G_{sol} the Gibbs free energy of the gas in solution. For a dense bubble layer with gas at low helium pressure and a planar configuration [189, 190]

$$\exp(\mu/kT - G_{sol}/kT) = \frac{N_{He}}{d_c}\left(\frac{2\pi\hbar^2}{mkT}\right)^{3/2}\left(\frac{kT}{\hbar\omega}\right)^3 \times \exp(-E_{sol}/kT) \tag{63}$$

with N_{He} the number of helium atoms in the bubble layer per unit area in m^{-2}, d_c the effective thickness of the bubble layer in m, m the mass of the ^{3}He atom in kg, and E_{sol} the helium solution energy as defined in Table 14 and Figure 44. Note that N_{He}/d_c represents the density of helium in the cavities. The effective layer thickness d_c equals 0.5 nm calculated as the product of the thickness of the cavity layer, 20 nm, and the void fraction of 2.5%. Substitution of

Eqs. (61)–(63) into Eq. (60) and integration with respect to time yields the following expression

$$\frac{N_{\mathrm{He}}}{N_0} = e^{-ft} \qquad \text{and}$$

$$f = \frac{\lambda^2 \omega}{12\pi L} \frac{c_{\mathrm{MgO}}}{d_c} \left(\frac{2\pi\hbar^2}{mkT}\right)^{3/2} \left(\frac{kT}{\hbar\omega}\right)^3 e^{-E_{\mathrm{perm}}/kT} \qquad (64)$$

The permeation energy, E_{perm}, is defined as $E_{\mathrm{perm}} = E_{\mathrm{sol}} + E_M$ and t is the time in s during which the permeation occurs. With a decrease in helium content of 86% ($N/N_0 = e^{-2}$) and the values given previously, the permeation energy can be calculated. This yields a permeation energy of 3.3 ± 0.3 eV, which agrees reasonably well with the dissociation energy of 3.6 eV for helium desorption from vacancy clusters calculated with the partial charge model.

As the predicted value is still somewhat higher than the experimental value, one has also to consider the mechanism of vacancy-assisted migration. The migration energy of a helium–vacancy complex is 2.3 eV calculated with the partial charge model. In order to obtain the activation energy for vacancy-assisted migration, this energy has to be added to the energy required to produce a thermal vacancy (a large concentration of vacancies is required for the process of vacancy-assisted migration). Thermal vacancies are produced at temperatures corresponding to an energy of 3.1 eV (half the Schottky defect energy as defined in Table 14) or more. Thus an activation energy of 5.4 eV for vacancy-assisted migration is obtained which makes this process much less feasible than helium permeation through the bulk.

6.4. Conclusions

The energetically most favorable configurations for vacancy clusters have been determined and three-dimensional growth of vacancy clusters is feasible, as this introduces less lattice polarization around the vacancy clusters. As a result, the energetically most favorable configuration of an eight-vacancy cluster is a cube. The partial charge model is more successful at predicting basic defect energies as the results calculated with the full charge model are in general too high when compared to experimental values reported in literature. The experimentally determined permeation energy of helium in MgO is 3.3 ± 0.3 eV as observed with NDP. This agrees reasonably well with the activation energy of 3.6 eV calculated with the partial charge model for helium dissociation from large vacancy clusters. However, the calculated activation energy for helium dissociation from a single oxygen vacancy is expected to be higher: 3.9 eV. This effect is due to lattice relaxation around the charged oxygen vacancy.

In particular, with respect to the helium defect interactions, more experimental data concerning helium desorption from MgO should be gathered in order to obtain a better understanding of the kinetics of the helium desorption process. This could be achieved with, for example, permeation experiments observed by NDP or by thermal helium desorption spectrometry (THDS). The predictions might be further improved by calculating the migration activation energies for helium from clusters explicitly. A THDS study has shown that in most cases helium release is characterized by the activation energy of $Q = 4.7 \pm 0.5$ eV with the maximum release temperature of $T_{\mathrm{max}} = 1830$ K. By applying a first order desorption model the pre-exponent factor is estimated as $\nu = 4.3 \times 10^{11}$ s^{-1}.

7. NANOCAVITIES IN SILICON

7.1. Introduction

In 1987 it was discovered by Griffioen et al. [191] that helium bubbles created by helium ion implantation in a silicon wafer could be turned into cavities by annealing to temperatures of about 800 °C. The helium and vacancies produced by implantation agglomerate in small bubbles at room temperature evolve further during annealing to larger bubbles from which at elevated temperature the helium is released. In contrast to metals, cavities in ceramics (e.g., silicon and the metal oxides considered here) are rather stable and the helium is trapped less strong. Therefore in most metals helium bubbles survive until temperatures close to melting. In silicon and most of the metal oxides (e.g., MgO, ZrO_2, UO_2) helium is released before the cavities disappear.

Subsurface layers of nanocavities are a good alternative for metal impurity gettering in silicon. Since Myers et al. [192] demonstrated in 1993 the use of the cavities for this purpose numerous studies (see, e.g., [193, 194]) have followed to unravel the details of formation and evolution of the cavities. Usually, the cavities are created by 30–40 keV helium ion implantation at depths of 200 nm but in this study the cavities are created at depths of 3–5.5 μm in order to leave more space to make certain devices in the remaining silicon top layer. Therefore, it is envisaged that the evolution of the bubbles and the release of helium is less influenced by the proximity of the surface. In other words, less of the reaction products evolving from dissociation of small helium vacancy complexes will reach the surface and therefore the bubble formation is more effective. The goal of the section is to understand the release of helium during the evolution of the defect layer from bubbles to cavities. In the first paper on release from cavities, release of 2.5 keV implanted helium ions could well be described by assuming a single cavity at shallow depth of 30 nm from the surface. Moreover, the description was restricted to cavities which were refilled with a small amount of helium (i.e., describing the low-pressure case). In this and many other studies release is monitored for bubbles at their initial pressure and for bubbles forming part of a bubble layer of thickness 10 to 50 bubble sizes [195, 196].

Thermal desorption spectrometry (TDS), NDP, non-Rutherford backscattering (NREBS), and NRA have been employed to measure the helium content and release rate during isothermal annealing at annealing temperatures of 800 and 1000 °C. TDS has also been used for isochronal annealing. X-TEM is used to monitor changes in morphology in the formed bubble layer. The helium release results can be modeled rather well when it is assumed that the helium initially is present in overpressurized bubbles. A single activation energy for helium release is revealed with a value of 1.83 (0.05) eV in agreement with the activation energy of helium permeation through silicon.

7.2. Experimental Methods

The Si(100) samples of a size of about 1 cm^2 have been implanted with the 3.5 MeV van de Graaf implanter at CERI Orleans. Two implantation energies, 1.00 and 1.55 MeV, respectively, have been used for the NDP analysis. ^{3}He was used as projectile. The amount of ^{3}He was determined by the ^{3}He$(n, p)^3$H reaction where a thermal neutron beam was used and the protons of the reaction with maximum kinetic energy of 574 keV were monitored by a Si detector. Protons originating from the helium bubble layer at a depth of 5.6 μm (1.55 MeV implantation) have 186 keV left when they leave the sample surface. Also for NRA the ^{3}He isotope is used in the ^{3}He$(d, p)^4$He reaction where the deuteron beam comes from the Orleans accelerator. In NREBS the enhanced scattering of MeV protons from ^{4}He is applied in the reaction ^{4}He$(p, p)^4$He.

It should be remarked that penetration depths and damage production of MeV ^{4}He and ^{3}He isotopes in silicon differ only marginally. Annealing for times varying up to 1.2×10^5 s were performed at two temperatures, 800 and 1000 °C. X-TEM pictures where taken at many stages of the annealing process. For one of the samples implanted at 1.6 MeV with 5×10^{16} cm^{-2} ^{4}He fluence, TDS measurements were performed on a small 1.5 mm^2 part of the sample. The sample was placed in a tungsten crucible that could be temperature programmed heated by electron bombardment. The helium release rate was monitored by a quadrupole mass spectrometer.

7.3. Annealing Results

Results of the remained fraction after several annealing steps are shown in Figure 49 for two temperatures and two fluences. Scattering of the results for NREBS is due to the limited sensitivity of the method with detection limit 3×10^{16} cm^{-2}. NDP has an intrinsic sensitivity 1000× better, but systematic errors are caused by the fact that the annealing treatments are not made for the same sample (in the 800 °C annealing series four samples were rotating) and reproducibility of the annealing oven. The NDP results showed that the bubble layer is located at a depth corresponding to the projected range of the helium. The distribution of the helium was approximated by a Gaussian at that depth. Results obtained for the FWHM of the Gaussian distribution before and after annealing are shown in Table 17. In all cases the first annealing step causes the distribution to narrow considerably. The value varies from 20 to 80% reduction; in particular for the low fluence the reduction is large. Thereafter, the width of the distribution does not change significantly. The 1000 °C annealing for the highest helium fluence yields some 20% wider distribution than the 800 °C annealing.

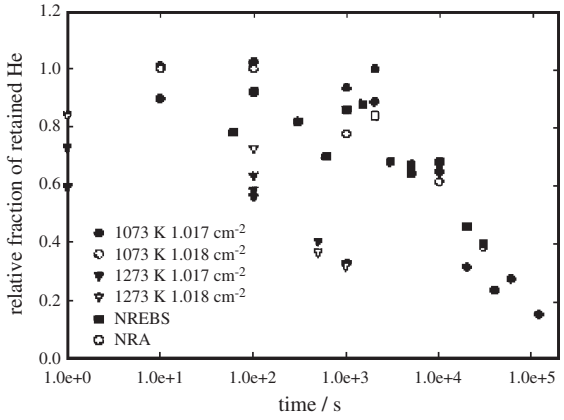


Figure 49. The retained helium fraction for silicon implanted at two fluences and 1.55 MeV helium ion energy for isothermal annealing at 1073 and 1273 K. NDP is used for helium detection. Results obtained with NREBS and NRA are also shown.

The two TEM pictures in Figure 50 show the cavity bands for 800 °C annealing of a 10^{17} cm^{-2} helium implanted sample for 10^3 and 6.1×10^4 seconds annealing time, respectively. The width of the layer with visible cavities has been determined to be 700 nm for 1000 s annealing and 500 nm for annealing up to 6.1×10^4 s. It has been found that bubbles generally increase in size by 20–50% for longer annealing periods. Note that the definition of the X-TEM width differs considerably from the previously defined FWHM determined by NDP. In X-TEM the outermost range is determined in which bubbles are visible, not taking into account the bubble density. Furthermore, the NDP measured width ignores the void fraction of the layer.

The *isothermal* helium release rate for a 5×10^{16} cm^{-2} ^{4}He implanted sample kept at 800 °C is shown in Figure 51. The advantage of TDS regarding release rate measurements is clearly demonstrated here. Due to the very high sensitivity of better than 10 ^{9}He s^{-1} and the fact that the rate rather than the content can be continuously measured the evolution of the release rate can be well monitored. It can be observed that the rate is not governed by a single relaxation time. The release process is about 10 times faster for short than for long annealing times.

Isochronal annealing with a heating rate in the temperature range from 300 to 1250 °C (600–1600 K) was not possible in one continuous run. Release rates would have turned the desorption spectrometer into saturation. Therefore, we adopted an annealing scheme where the isochronal annealing starts first with 800 K as maximum temperature and it is repeated every time to a 50 K higher temperature. The method, which is called partial annealing, keeps the release rates low and has, as will be shown later, certain advantages for desorption mechanism analysis. The results are shown in Figure 52. The plots show the desorption rates during both the heating and the cooling (thin line). With this annealing scheme the maximum release rate is found at temperatures

Table 17. FWHM of the helium distributions measured by NDP before and after annealing.

Energy (MeV)	Fluence (10^{17} cm^{-2})	Annealing temperature (°C)	As implanted (μm)	Annealed (μm)
1.55	1	800	0.37	0.31
1.55	0.5	800	0.38	0.23
1.00	0.5	800	0.27	0.11
1.55	1	1000	0.47	0.37
1.55	0.5	1000	0.37	0.25
1.00	0.5	1000	0.27	0.12

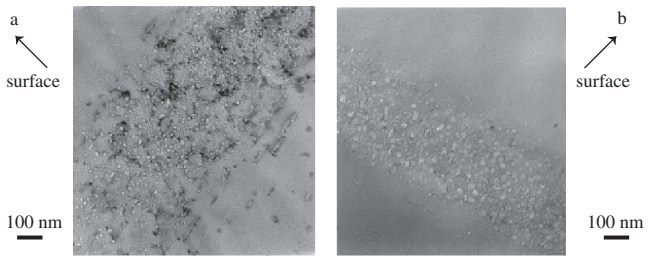


Figure 50. Micrographs (X-TEM) for helium implanted silicon (1×10^{17} cm^{-2}; 1.55 MeV) after annealing at 1073 K for 1000 (a) and 6×10^4 s (b), respectively.

higher than 1300 K (1000 °C). A small remaining fraction of helium is not released until close to the melting point of silicon. The spike is due to bursting of blisters, which develop in the final stage of the annealing process.

7.4. Bubble Evolution and Helium Release

7.4.1. First Stages of the Annealing Process

The results indicate that in the first steps of the process, in the case of isothermal annealing, during heating to the annealing temperature, a complicated defect evolution takes place. Small bubbles or platelets [197] transform into larger bubbles by loop punching and later by Ostwald ripening. It is very likely that all the gas remains contained in the bubble layer during this coarsening process. At low fluence the initial concentration of the helium in the layer is lower than for high fluence so that the layer thickness shrinks more during annealing. The bubbles will be overpressurized. Thermal growth of cavities by interstitial emission or vacancy capture is slow at 800 °C and even at 1000 °C it will take considerable time. It is thought that the cavities move by surface diffusion and eventually coalesce. Equilibrium pressures for bubbles with a size of 7 nm are of the order of 1 GPa. Thus in modeling the gas release this high pressure should be taken into account.

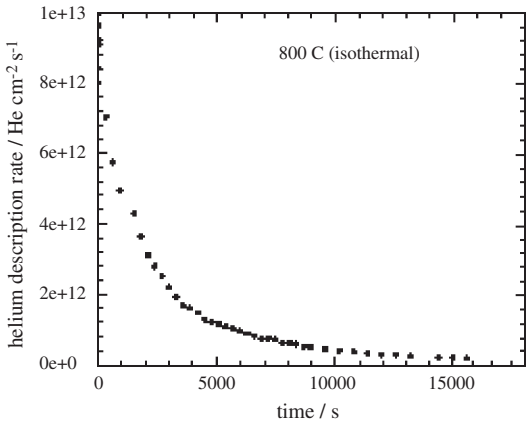


Figure 51. Release rate during isothermal annealing of 5×10^{16} cm^{-2} He implanted Si measured by TDS.

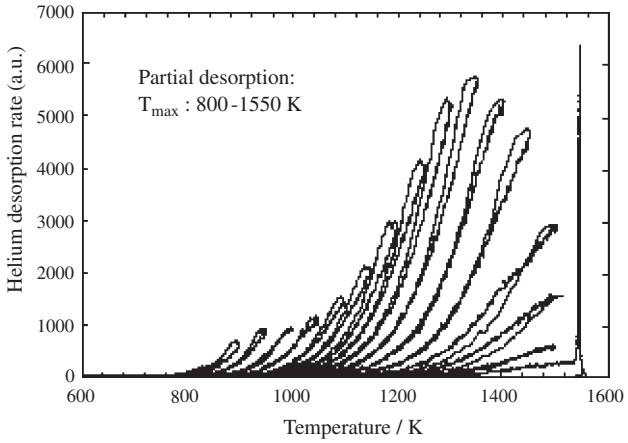


Figure 52. Partial desorption results for small samples similarly implanted as the sample in Figure 51. The heating rate was 3.6 K s^{-1}.

7.4.2. Helium Release Models for Release from Bubbles and Bubble Agglomerates

A good approximation of the helium release from a gas filled cavity is to assume that the gas release is due to a quasi-stationary flow of gas which is sustained by the gradient that exists from the concentration of dissolved gas near the cavity surface equilibrium to the zero concentration near the external surface. First, can be easily calculated that in our experiments time constants to establish equilibrium between gas in the cavity and the dissolved gas near the cavity surface are negligible compared to the observed times required to release a substantial amount of gas. Second, the relaxation time required to establish a stationary flow $\tau_D = L^2/2D$ where L is the distance from cavity to surface and D the helium diffusivity [198] is also small and of the order of seconds. Thus, the gas release rate can be obtained by solving the stationary diffusion equation for the cavity geometry of interest with the appropriate equilibrium concentration of gas c_{eq} as boundary condition near the cavity and zero concentration at the external surface. The concentration is given by equations similar to those in Section 6 for He in MgO, that is, Eqs. (62) and (63) with $\omega = 2\pi\nu$ with $\nu = 4 \times 10^{13}$ s^{-1} for He in Si [191, 192]. The gas flow is found by integrating the diffusion current $J_D = D(dc/dx)_{x=0}$ over the surface area.

For *isolated bubbles* with radius r at distance L from the surface the gas flow

$$\frac{dN}{dt} = \frac{3NP}{rL} \tag{65}$$

with P the permeability, for example, measured by [199] and applied by Griffioen et al. to describe release of low-pressure helium refilled cavities. The description has shown to work satisfactorily also for release from cavities in SPE epitaxial layers [200].

For a *bubble layer with gas at low helium pressure* a similar release rate can be derived, but because of the different geometry the flow is given by

$$\frac{dN}{dt} = \frac{NP}{d_c L} \tag{66}$$

with d_c the effective thickness of the cavity layer (i.e., void fraction times layer thickness).

It can be calculated that interbubble diffusion of helium in the bubble layer is fast compared with the release rate constant. Therefore, it is justified to consider the cavity layer as a united gas reservoir. For nonequilibrium bubbles at high helium pressure the concentration can no longer be derived based on the ideal gas law. Then, the high-pressure expression of the chemical potential μ has to be used.

In Figure 53 the concentration of dissolved gas is calculated according to the experimentally determined EOS of Mills et al. [201] for pressures up to 3 GPa. It is clearly observed that for high pressures (high gas concentrations) in the cavities the solubility strongly increases. Using the EOS of Mills et al. and assuming an initial pressure of 2 GPa (20 kbar) the retention of helium gas is calculated for the experiments shown in Figure 51. The results of the calculations shown in Figure 54a describe the experimental results for the 1073 K annealing rather well (see Fig. 51) when an activation energy for permeation of 1.8 eV and an effective thickness of 200 nm are assumed. The results for 1273 K (not shown) fit well to the experimental result. The value assumed for the effective thickness has not yet been experimentally verified. It is very difficult to derive the void fraction from the TEM pictures. Note that no attempt was made to fit the data. The results are the same when the EOS proposed by Trinkaus is used [202]. In Figure 54b the calculated isothermal release rate is plotted for the 800 °C isothermal annealing which, in line with the experimental results, shows a fast release rate at the beginning of the release process.

7.4.3. Isochronal Annealing: Activation Energy

The partial annealing experiments can give extra information on the release mechanism by considering the initial part of each desorption curve. In principle, the release is a first-order desorption process

$$\frac{dN}{dt} = -Nf \tag{67}$$

with $f = f_o \exp(-E/kT)$ the temperature dependent release frequency and N the number of retained He. Considering

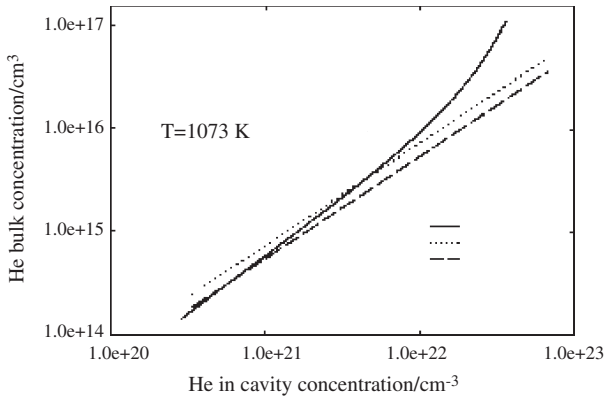


Figure 53. Helium solubility versus Helium density in cavities from the Mills et al. EOS [···] (—), ideal gas law (·····), and [191, 192], respectively.

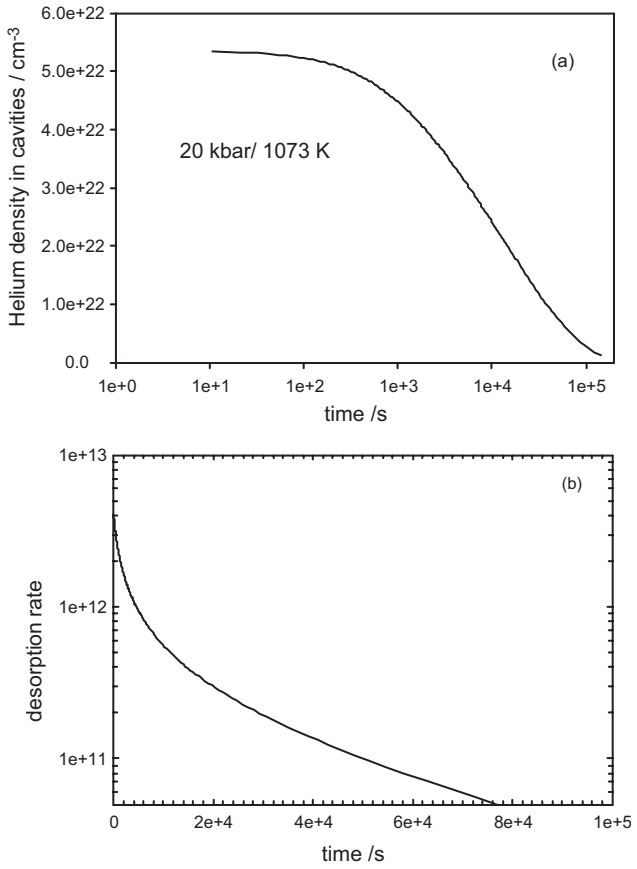


Figure 54. Calculated retained fraction (a) and release rate (b) during annealing for a 1.55 MeV helium implanted Si sample (dose 5×10^{16} cm^{-2}). The Mills et al. equation of state has been used in the theory described in the text.

only the first part of the release N can be considered to be constant and plotting of the release rate vs $1/T$ will show Arrhenius behavior. This is done in Figure 55, and it is found that the energy E derived from the slopes of the partial annealing curves is nearly constant over the whole temperature range. We found $E = 1.83$ (0.05) eV for the activation energy of the process. From the measurements also the different pre-exponential factors $N_i f_{o,i}$ of the different partial annealing treatments could be derived with i denoting the ith partial annealing. Since we know N_i we can calculate the factors $f_{o,i}$. These values have been plotted in Figure 56 vs the helium amount in the cavities. Again it is seen that the prefactor is higher when the concentration is still high. At very low retained fractions the prefactor is reduced even further. This we might ascribe to the observed enlargement of the cavities at very long annealing times. A possible mechanism might be *pressure induced thermal creep* because there is a considerable pressure exerted on the layer by the gas containing bubbles.

In another study Grisolia et al. report in more detail on the growth mechanism of bubbles and cavities in the cavity band. Variation of annealing temperature from 700 to 900 °C leads to increased coarsening. In Figure 57 it is observed that at 900 °C a majority of large cavities of size 20–50 nm survive which are concentrated in a

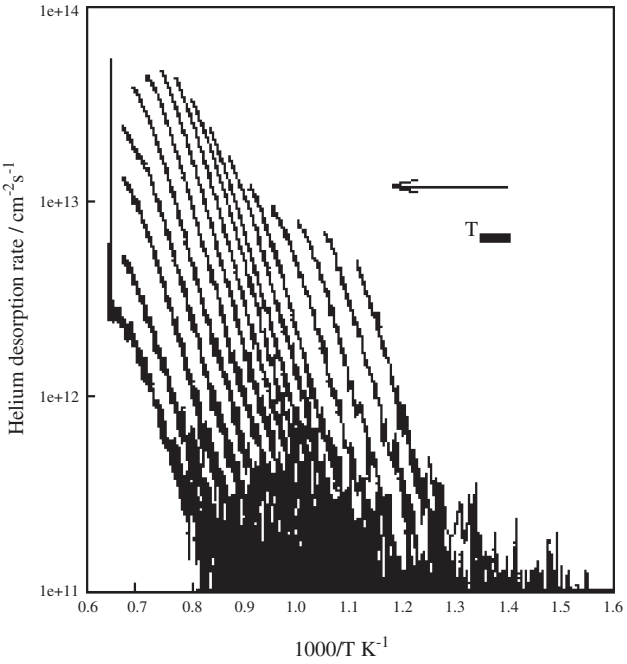


Figure 55. Partial desorption Arrhenius plot for TDS results shown in Figure 52.

narrow band below the surface. The cavity density and the size have been plotted in Figure 58. It shows a considerable decrease of density corresponding with an increase of size at temperatures beyond 800 °C. It was derived that the cavity growth was governed by an activation energy of 1.65 eV. Oliviero et al. studied the influence of the implantation temperature on the density and morphology of the bubbles and the cavities. Typical TDS results in Figure 59 for 50 keV implanted helium at low implantation temperature (473 K) show release in release peaks at temperatures around 800 and around 1200 K, respectively. The low-temperature release occurs when helium–vacancy complexes formed during the implantation break up and form larger complexes. A part of the helium is not retrapped in the newly formed larger complexes and escapes through the surface. The high-temperature release is by helium

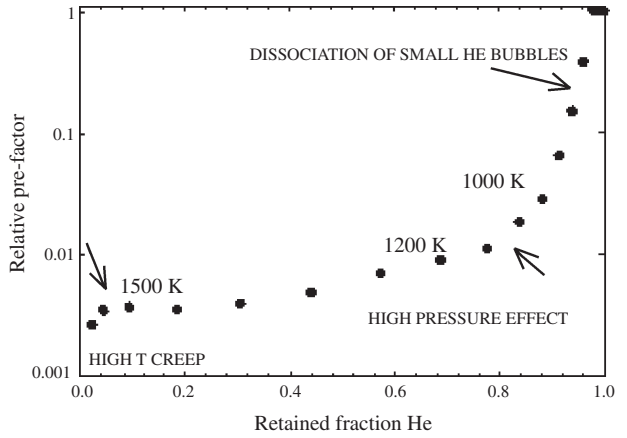


Figure 56. Relative prefactor derived from Figure 55.

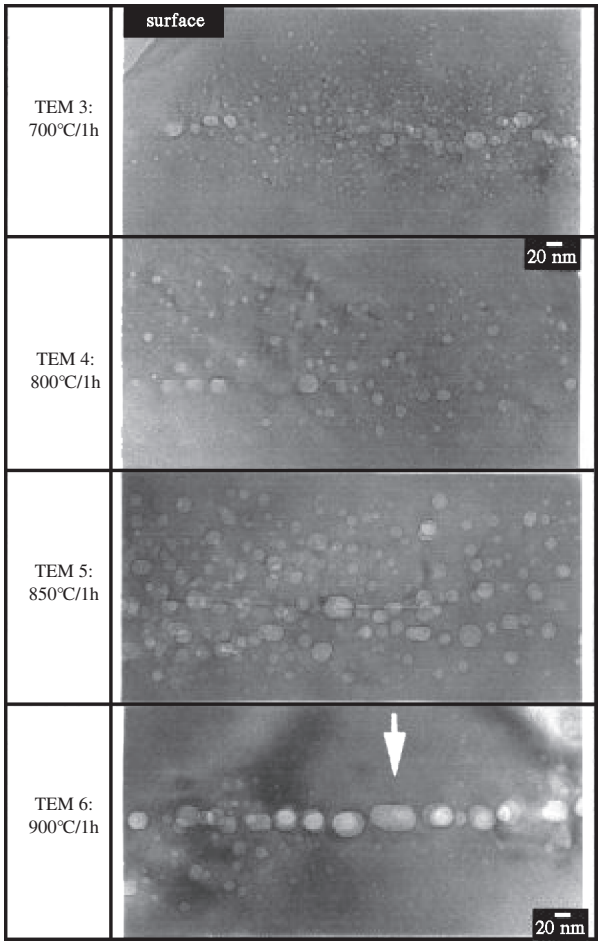


Figure 57. Micrographs showing annealing of helium implanted silicon at temperature varied from 700 to 900 °C.

permeation from bubbles when they have formed. The results are qualitatively similar to the results earlier shown for the 1.5 MeV helium implanted samples. At increasingly higher temperature (673 and 773 K) the first release peaks change their shape and the temperature of maximum helium release. This indicates that depending on the implantation temperature a certain size of helium

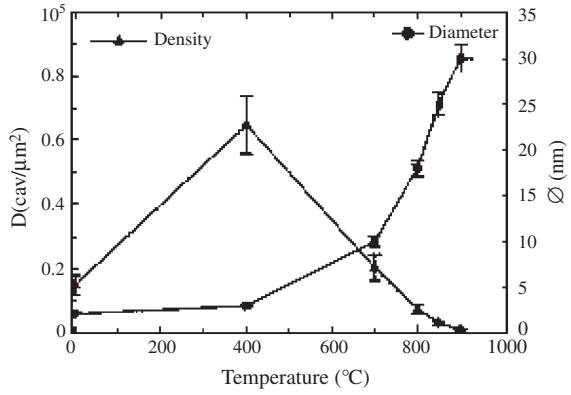


Figure 58. Cavity density and size vs temperature for helium implanted silicon.

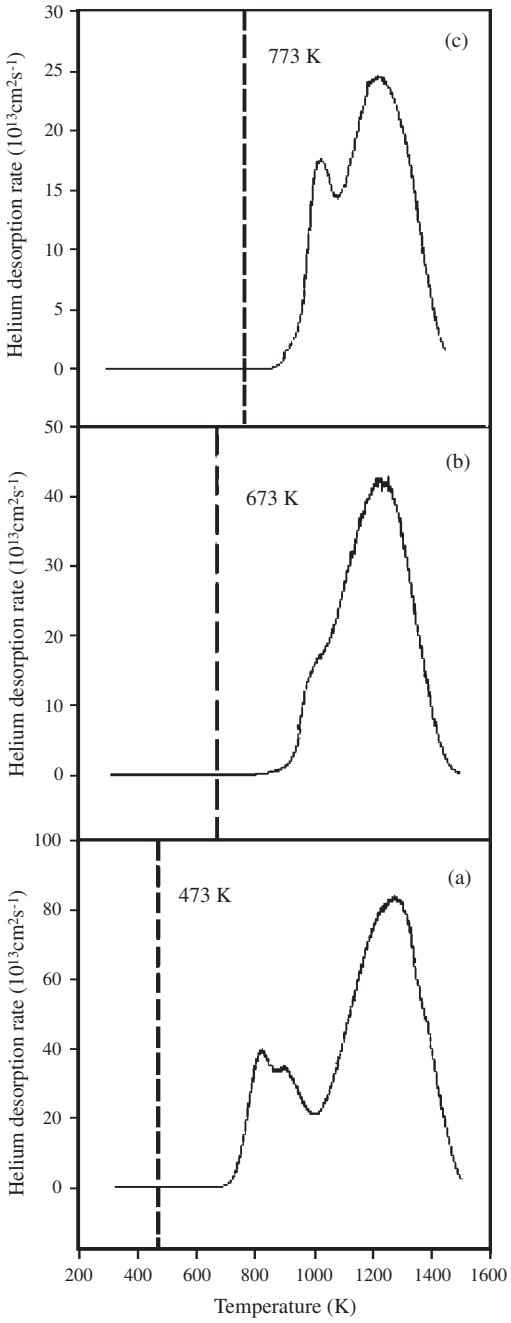


Figure 59. Helium desorption spectrum obtained after helium implantation at (a) 473, (b) 673, and (c) 773 K in silicon (5×10^{16} cm^{-2}, 50 keV). Heating rate: 5 K/s. The dashed line indicates the temperature of implantation.

vacancy complex can survive. Upon further annealing release from these complexes is observed. For implantation at 875 K (desorption spectra are not shown) the helium release is very small, because very few bubbles and cavities are formed. In Figure 60 the retained helium and the bubble/void morphology are shown for the different helium implantation temperatures. At low T small bubbles are observed and at 873 K when all helium is released larger cavities remain. At 1075 K implantation no bubbles are formed because no stable helium

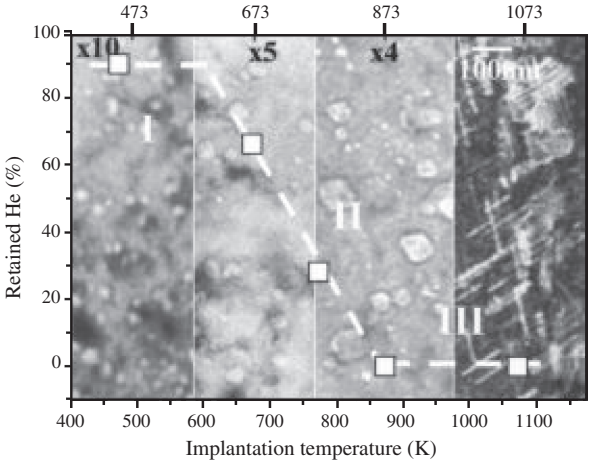


Figure 60. Helium retained in the silicon sample after helium implantation (5×10^{16} cm^{-2}, 50 keV) vs the implantation temperature. In the background, typical TEM images are presented in bright-field underfocus conditions for the 473, 673, and 873 K implantations and in weak-beam, dark-field conditions (g, $7g$) for the 1073 K implantation.

vacancy clusters can be nucleated. Instead, self-interstitial agglomerates have survived which are arranged in ⟨311⟩ defects and dislocation arrays.

In summary, the first phase of helium release from helium implanted silicon is ascribed to dissociation of small helium vacancy complexes, and in the second phase of release it is ascribed to helium permeating from a helium bubble layer to the sample surface. The fact that bubbles are overpressurized and that release does not occur from single bubbles but from a bubble layer should be taken into account. In the final stage of helium release bubble volume might increase by thermal creep thus slowing down the permeation.

8. FUTURE PROSPECTS

In the future the research on nanoprecipitates and nanocavities will develop in several interesting directions. For the nanoprecipitates better definition of the size of the implantation induced nanprecipitates or nanoclusters can be foreseen by combining ion implantation and ion lithographic mask technology. When ion implantation can be confined to well defined open areas in a mask the number of atoms implanted per opening in the mask is well defined and therefore the number of atoms constituting the nanoparticle that is formed after thermal annealing. In that manner monodisperse particles can be produced. The number of possible combinations that can be made of implanted atoms is virtually inexhaustible. Thus exotic compounds can be produced which will not form without the pressure exerted on the particle by the host matrix. Furthermore, the host matrix will impose a certain preference for interfaces and phases of the embedded particles. Also nanocavities in the metal oxides can be used as a template to create hollow particles when the internal surfaces allow wetting by implanted foreign atoms. For nanocavities in silicon a natural step in the development is to develop patterned nanostructures of cavities in ultra-large-scale integrated devices so that nanocavities for

impurity gettering are, on purpose, located in the proximity of electrically active areas of the elements of the device.

GLOSSARY

Auger process A non-radiative process in which an atom in an excited state undergoes a transition to a lower state by the emission of a bound electron (Auger electron) rather than by the emission of an X-ray. Auger electrons have a characteristic energy, unique to each element from which it was emitted. These electrons are collected and sorted according to energy to give compositional information about the specimen. Since Auger Electrons have relatively low energy they are only emitted from the bulk specimen from a depth of smaller than 3 nm.

Contrast transfer function Describes the imaging properties of the objective lens. It is independent of the object and is related to the sine function of the phase introduced by the aberration and defocus.

Doppler effect Refers to the change in the wavelength (or frequency) of energy in the form of waves, e.g., sound or light, as a result of motion of either the source or the receiver of the waves; the effect is named for the Austrian scientist Christian Doppler, who was professor at the Polytechnic in Prague and who demonstrated the effect for sound. If the source of the waves and the receiver are approaching each other, the frequency of the waves will increase and the wavelength will be shortened—sounds will become higher pitched and light will appear bluer. If the sender and receiver are moving apart, sounds will become lower pitched and light will appear redder.

Neutron An uncharged elementary particle of slightly greater mass than the proton. James Chadwick discovered it in 1932. The neutron, like other particles, also possesses certain wave properties and the field of neutron optics is concerned with such topics as the neutron diffraction and polarization of beams of neutrons. James Chadwick received the Nobel Prize in physics in 1935.

Positron annihilation A result of an encounter of the electron with its antiparticle-positron. The energy released by the annihilation forms two highly energetic gamma photons, which travel in opposite direction. These gamma rays provide a useful analysis tool, which has found many practical applications in physics and chemistry.

Positron An elementary particle having the same mass and spin as an electron but having a positive charge equal in magnitude to that of the electron's negative charge; Actually it can be regarded as the antiparticle of the electron and it is also called anti-electron and positive electron. Dirac has postulated the positron in 1930 as the "negative" energy extension of his theory of electron energy levels. Dirac received the Nobel Prize in physics in 1933.

Positronium A short-lived atomic system consisting of a positron and an electron bound together. Positronium is a temporary state in the decay of the positron. A positron in free space will quickly interact with an electron via the Coulomb force, and the positron and electron will for a short time form a bound state. This state is considered to be an atom in the sense that it is a hydrogen atom with the proton replaced by a positron.

Rutherford back-scattering technique An analytical tool that uses elastic scattering of 0.1–3 MV charged particles to analyze the surface and the outer few micrometers of solids. The depth distribution of constituents can be reconstructed with a depth resolution of 10–20 nm. The RBS technique is non-destructive since the erosion and the radiation degradation of the sample material by the particle impact is negligible. Rutherford was a British physicist who first explained that radioactivity is produced by the disintegration of atoms and discovered that alpha particles consist of helium nuclei. For these achievements, Rutherford was awarded the 1908 Nobel Prize in chemistry.

Transmission electron microscopy Because the wavelength of electrons is much smaller than that of light (of the order of pico-meters), the optimal resolution attainable for TEM images is many orders of magnitude better than that from a light microscope. Thus, TEMs can reveal the finest details of internal structure—in some cases as small as individual atoms. The point-point-resolution depends not only on the wavelength but also on the spherical aberration coefficient of the objective lens. Under special circumstances, atoms can be imaged at magnifications greater than 15 million times. Materials for TEM must be specially prepared to thicknesses, which allow electrons to transmit through the sample. Max Knoll and Ernst Ruska in Germany developed it in 1931. Ernst Ruska received the Nobel Prize in physics in 1986.

Weak phase object Image formation in a high resolution electron microscope is an interference phenomenon. The transmission of the incident wave by the object (diffraction) introduces phase shifts to the diffracted beams. For thin crystals, also called weak phase objects, the phase is assumed to be the constant. Frits Zernike at the University of Groningen, The Netherlands, discovered the phase contrast phenomenon in 1930. He received the Nobel Prize in physics in 1953.

ACKNOWLEDGMENTS

Many thanks are due to our collaborators and (former) Ph.D. students for their work, contributions, and discussions: M. A. van Huis, H. Schut, C. V. Falub, P. E. Mijnarends, S. W. H. Eijt, A. V. Fedorov, S. W. H. Eijt, B. J. Kooi, R. L. Zimmerman, F. Labohm, E. A. C. Neeft, R. J. M. Konings, J. de Roode, T. Heijenga, J. Kuriplach, J. de Vries, R. A. Hakvoort, M. R. Ijpma, C. C. Griffioen, J. H. Evans, S. Godey, E. Ntsoenzok, T. Sauvage, M. F. Beaufort, J. F. Barbot, P. C. de Jong, A. H. Reader, D. J. Gravesteijn, and A. A. van Gorkum. The work has been supported by the Foundation for Fundamental Research on Matter (FOM-Utrecht, The Netherlands) and has been made possible by financial support from the Netherlands Organization for Research (NWO, The Hague, The Netherlands).

REFERENCES

1. R. D. Miller, *Science* 286, 421 (1999).
2. A. Polman, *Nature Mater.* 1, 10 (2002).
3. P. J. Schultz and K. G. Lynn, *Rev. Mod. Phys.* 60, 701 (1988).

4. L. C. Feldman and J. W. Mayer, "Fundamentals of Surface and Thin Film Analysis." Elsevier, Amsterdam, 1986.
5. A. Dupasquier and G. Ottaviani, in "Positron Spectroscopy of Solids" (A. Dupasquier and A. P. Mills, Jr., Ed.), pp. 581–657. IOS, Amsterdam, 1995.
6. A. van Veen, H. Schut, and P. E. Mijnarends, in "Positron Beams and their Applications" (P. G. Coleman, Ed.), p. 191. World Scientific, Singapore, 2000.
7. "Positron Beams and Their Applications" (P. G. Coleman, Ed.). World Scientific, Singapore, 2000.
8. R. H. Howell, in "Positron Beams and Their Applications" (P. G. Coleman, Ed.), p. 307. World Scientific, Singapore, 2000.
9. J. P. Peng, K. G. Lynn, P. Asoka-Kumar, and D. P. Becker, *Phys. Rev. Lett.* 76, 2157 (1996).
10. A. van Veen, F. Labohm, H. Schut, J. de Roode, T. Heijenga, and P. E. Mijnarends, *Appl. Surf. Sci.* 116, 39 (1997).
11. A. van Veen, H. Schut, J. de Roode, F. Labohm, C. V. Falub, S. W. H. Eijt, and P. E. Mijnarends, *Mater. Sci. Forum* 363–365, 415 (2001).
12. S. W. H. Eijt, C. V. Falub, P. E. Mijnarends, and A. van Veen, *Mater. Sci. Forum* 363–365, 588 (2001).
13. S. W. H. Eijt, C. V. Falub, A. van Veen, H. Schut, P. E. Mijnarends, M. A. van Huis, and A. V. Fedorov, *Mater. Res. Soc. Symp. Proc.* 647, 014.11 (2001).
14. A. van Veen, M. A. van Huis, A. V. Fedorov, H. Schut, C. V. Falub, S. W. H. Eijt, F. Labohm, B. J. Kooi, and J. Th. M. de Hosson, *Mater. Res. Soc. Symp. Proc.* 647, 011.24 (2001).
15. M. A. van Huis, A. van Veen, H. Schut, C. V. Falub, S. W. H. Eijt, P. E. Mijnarends, and J. Kuriplach, *Phys. Rev. B* 65, 085416 (2002).
16. C. V. Falub, P. E. Mijnarends, S. W. H. Eijt, M. van Huis, A. van Veen, and H. Schut, *Nucl. Instrum. Methods B* 191, 275 (2002).
17. S. W. H. Eijt, C. V. Falub, A. van Veen, H. Schut, and P. E. Mijnarends, *Appl. Surf. Sci.* 194, 234 (2002).
18. C. V. Falub, P. E. Mijnarends, S. W. H. Eijt, M. van Huis, A. van Veen, and H. Schut, *Phys. Rev. B* 66, 0754xx (2002).
19. P. A. M. Dirac, *Proc. Roy. Soc. London Ser. A* 126, 360 (1930).
20. P. A. M. Dirac, *Proc. Cambridge Philos. Soc.* 26, 361 (1930).
21. C. D. Anderson, *Science* 76, 238 (1932); *Phys. Rev.* 41, 405 (1932); *Phys. Rev.* 43, 491 (1933).
22. A. Ore and J. L. Powell, *Phys. Rev.* 75, 1696 (1949).
23. P. Kubica and A. T. Stewart, *Phys. Rev. B* 4, 2490 (1975).
24. J. L. Lee, J. T. Waber, Y. K. Park, and J. Th. M. De Hosson, *J. Mater. Sci. Eng.* 81, 379 (1986).
25. S. Mohorovicic, *Astron. Nachr.* 253, 94 (1934).
26. A. E. Ruark, *Phys. Rev.* 68, 278 (1945).
27. M. Deutsch, *Phys. Rev.* 82, 455 (1951).
28. O. Halpern, *Phys. Rev.* 94, 904 (1954).
29. A. Dupasquier, Positron solid state physics, in "Proc. of the International School of Physics 'Enrico Fermi'" (W. Brandt and A. Dupasquier, Eds.), Course LXXXIII, p. 510. IOS Press, Amsterdam, 1983.
30. T. Hyodo, Positron spectroscopy of solids, in "Proc. of the International School of Physics 'Enrico Fermi'" (W. Brandt and A. Dupasquier, Eds.), Course CXXV, p. 419. IOS Press, Amsterdam, 1995.
31. M. Deutsch, "Adventures in Experimental Physics," Vol. 4, p. 64. Princeton Univ. Press, Princeton, 1975.
32. S. Berko and H. N. Pendleton, *Annu. Rev. Nucl. Sci.* 30, 543 (1980).
33. K. G. Lynn, Positron solid state physics, in "Proc. of the International School of Physics 'Enrico Fermi'" (W. Brandt and A. Dupasquier, Eds.), Course LXXXIII, p. 609. IOS Press, Amsterdam, 1983.
34. O. E. Mogensen, *J. Chem. Phys.* 60, 998 (1974).
35. K. G. Lynn, A. P. Mills, Jr., R. N. West, S. Berko, K. F. Canter, and L. O. Roellig, *Phys. Rev. Lett.* 54, 1702 (1985).
36. D. M. Chen, S. Berko, K. F. Canter, K. G. Lynn, A. P. Mills, Jr., L. O. Roellig, P. Sferlazzo, M. Weinert, and R. N. West, *Phys. Rev. B* 39, 3966 (1989).
37. A. H. Weiss, in "Positron Beams and their Applications" (P. G. Coleman, Ed.), p. 129. World Scientific, Singapore, 2000.
38. I. K. MacKenzie, *Phys. Lett. A* 30, 115 (1969).
39. S. Berko and J. S. Plaskett, *Phys. Rev.* 112, 1877 (1958).
40. S. Berko, M. Haghgooie, and J. J. Mader, *Phys. Lett. A* 63A, 335 (1977).
41. R. N. West, Positron spectroscopy of solids, in "Proc. of the International School of Physics 'Enrico Fermi'" (A. Dupasquier and A. P. Mills, Jr., Eds.), Course CXXV, p. 75. IOS Press, Amsterdam, 1995.
42. K. G. Lynn, A. P. Mills, R. N. West, S. Berko, K. F. Canter, and L. O. Roellig, *Phys. Rev. Lett.* 54, 1702 (1985).
43. D. M. Chen, S. Berko, K. F. Canter, K. G. Lynn, A. P. Mills, L. O. Roellig, P. Sferlazzo, M. Weinert, and R. N. West, *Phys. Rev. Lett.* 58, 921 (1987).
44. D. M. Chen, S. Berko, K. F. Canter, K. G. Lynn, A. P. Mills, L. O. Roellig, P. Sferlazzo, M. Weinert, and R. N. West, *Phys. Rev. B* 39, 3966 (1989).
45. R. H. Howell, P. Meyer, I. J. Rosenberg, and M. J. Fluss, *Phys. Rev. Lett.* 54, 1698 (1985).
46. I. K. MacKenzie, J. A. Cadly, and R. R. Gingerich, *Phys. Lett. A* 33, 279 (1970).
47. I. K. MacKenzie, Positron solid state physics, in "Proc. of the International School of Physics 'Enrico Fermi'" (W. Brandt and A. Dupasquier, Eds.), Course LXXXIII, p. 75. IOS Press, Amsterdam, 1983.
48. A. F. Makhov, *Sov. Phys. Solid State* 2, 1934 (1961).
49. A. van Veen, H. Schut, J. de Vries, R. A. Hakvoort, and M. R. Ijpma, *AIP Conf. Proc.* 218, 171 (1990).
50. K. G. Lynn, J. R. MacDonald, R. A. Boie, L. C. Feldman, J. D. Gabbe, M. F. Robbins, E. Bonderup, and J. Golovchenko, *Phys. Rev. Lett.* 38, 241 (1977).
51. K. G. Lynn, W. E. Frieze, and P. J. Schultz, *Phys. Rev. Lett.* 52, 1137 (1984).
52. A. P. Mills, L. Pfeiffer, and P. M. Platzman, *Phys. Rev. Lett.* 51, 1085 (1983).
53. R. H. Howell, M. J. Fluss, I. J. Rosenberg, and P. Meyer, *Nucl. Instrum. Methods B* 10/11, 373 (1985).
54. D. Schödlbauer, P. Sperr, G. Kögel, and W. Triftshäuser, *Nucl. Instrum. Methods B* 34, 258 (1988).
55. P. Willutzki, J. Störmer, G. Kögel, P. Sperr, D. T. Britton, R. Steindl, and W. Triftshäuser, *Mater. Sci. Forum* 175–178, 237 (1995).
56. R. Suzuki, T. Mikado, H. Ohgaki, M. Chiwaki, T. Yamazaki, and Y. Kobayashi, *AIP Conf. Proc.* 303, 526 (1994).
57. P. Willutzki, J. Störmer, G. Kögel, P. Sperr, D. T. Britton, R. Steindl, and W. Triftshäuser, *Mater. Sci. Forum* 175–178, 237 (1995).
58. R. Suzuki, T. Mikado, H. Ohgaki, M. Chiwaki, T. Yamazaki, and Y. Kobayashi, *AIP Conf. Proc.* 303, 526 (1994).
59. K. G. Lynn, L. O. Roellig, A. P. Mills, and A. Moodenbaugh, in "Atomic Physics with Positrons" (J. W. Humberston and E. A. G. Armour, Eds.), p. 161. Plenum Press, New York, 1987.
60. Y. Itoh, Z. Peng, K. Lee, M. Ishii, A. Goto, N. Nakanishi, M. Kase, and Y. Ito, *Appl. Surf. Sci.* 116, 68 (1997).
61. R. H. Howell, R. A. Alvarez, and M. Stanek, *Appl. Phys. Lett.* 40, 751 (1982).
62. C. Hugenschmidt, G. Kögel, K. Schreckenbach, P. Sperr. M. Springer, B. Straßer, and W. Triftshäuser, *Appl. Surf. Sci.* 149, 7 (1999).
63. F. Zernike, *Z. Tech. Phys.* 16, 454 (1935); *Phys. Z.* 36, 848 (1935).
64. J. Th. M. De Hosson, in "Handbook of Microscopy" (S. Amelinckx, D. van Dyck, J. van Landuyt, and G. van Tenderloo, Eds.), Vol. 3, pp. 5–110. VCH, Weinheim, 1997.

65. P. Buseck, J. Cowley, and L. Eyring, "High Resolution Transmission Electron Microscopy." Oxford Univ. Press, London, 1992.
66. J. C. H. Spence, "Experimental High-Resolution Electron Microscopy," 2nd ed. Oxford Univ. Press, London, 1988.
67. D. B. Williams and C. B. Carter, "Transmission Electron Microscopy, A Textbook for Materials Science." Plenum, New York, 1996.
68. J. M. Cowley and A. F. Moodie, *Acta Crystallogr.* 10, 609 (1957).
69. M. O'Keefe, *Ultramicroscopy* 47, 282 (1992).
70. O. Scherzer, *J. Appl. Phys.* 20, 20 (1949).
71. M. Haider, H. Rose, S. Uhlemann, E. Schwan, B. Kabius, and K. Urban, *Ultramicroscopy* 75, 53 (1998).
72. P. Chakraborty, *J. Mater. Sci.* 33, 2235 (1998).
73. C. W. White, J. D. Budai, S. P. Withrow, J. G. Zhu, E. Sonder, R. A. Zuhr, A. Meldrum, D. M. Hembree, Jr., D. O. Henderson, and S. Prawer, *Nucl. Instrum. Methods B* 141, 228 (1998).
74. R. L. Zimmerman, D. Ila, E. K. Williams, S. Sarkisov, D. B. Poker, and D. K. Hensley, *Nucl. Instrum. Methods B* 141, 308 (1998).
75. D. Ila, Z. Wu, C. C. Smith, D. B. Poker, D. K. Hensley, C. Klatt, and S. Kalbitzer, *Nucl. Instrum. Methods B* 127/128, 570 (1997).
76. A. van Veen, H. Schut, A. V. Fedorov, F. Labohm, E. A. C. Neeft, and R. J. M. Konings, *Nucl. Instrum. Methods B* 148, 768 (1999).
77. P. J. Schulz and K. G. Lynn, *Rev. Mod. Phys.* 60, 701 (1988).
78. A. V. Fedorov, M. A. van Huis, A. van Veen, and H. Schut, *Nuclear Instrum. Methods B* 166–167, 215 (2000).
79. B. J. Kooi, A. van Veen, J. Th. M. de Hosson, H. Schut, A. V. Fedorov, and F. Labohm, *Appl. Phys. Lett.* 76, 1110 (2000).
80. M. A. van Huis, A. V. Fedorov, A. van Veen, C. V. Falub, S. W. H. Eijt, B. J. Kooi, J. Th. M. de Hosson, T. Hibma, and R. L. Zimmerman, *Nucl. Instrum. Methods B* 191, 442 (2002).
81. J. F. Ziegler, J. P. Biersack, and V. Littmark, "The Stopping and Range of Ions in Solids" (J. F. Ziegler, Eds.). Pergamon Press, New York, 1985.
82. R. L. Zimmerman, D. Ila, E. K. Williams, D. B. Poker, D. K. Hensley, C. Klatt, and S. Kalbitzer, *Nucl. Instrum. Methods B* 148, 1064 (1999), and references therein.
83. W. T. Doyle, *Phys. Rev.* 111, 1067 (1958).
84. J. I. Langford and A. J. C. Wilson, *J. Appl. Cryst.* 11, 102 (1978).
85. C. Kittel, "Introduction to Solid State Physics," 6th ed., p. 134. Wiley, New York, 1986.
86. U. Kreibig, A. Hilger, H. Hövel, and M. Quinten, in "Large Clusters of Atoms and Molecules" (T. P. Martin, Ed.), pp. 475–493. Kluwer Academic, The Netherlands, 1996.
87. D. L. Mills, "Nonlinear Optics." Springer, Berlin, 1998.
88. M. Backhaus-Ricoult and S. Laurent, *Mater. Sci. Forum* 294–296, 173 (1999).
89. "Metal Clusters" (W. Ekardt, Ed.). Wiley, New York, 1999.
90. "Semiconductor Nanoclusters" (P. V. Kamat and D. Meisel, Eds.). Elsevier, New York, 1997.
91. D. Davidovic and M. Tinkham, *Phys. Rev. Lett.* 83, 1644 (1999).
92. A. P. Alivisatos, *Science* 271, 933 (1996).
93. "Positron Spectroscopy of Solids" (A. Dupasquier and A. P. Mills, Jr., Eds.). IOS Press, Amsterdam, 1995.
94. M. J. Puska and R. M. Nieminen, *Rev. Mod. Phys.* 66, 841 (1994).
95. R. N. West, in [8], p. 75.
96. S. S. Rajput, R. Prasad, R. M. Singru, W. Trifthäuser, A. Eckert, G. Kögel, S. Kaprzyk, and A. Bansil, *J. Phys.: Condens. Matter* 5, 6419 (1993).
97. Y. Nagai, M. Hasegawa, Z. Tang, A. Hempel, K. Yubuta, T. Shimamura, Y. Kawazoe, A. Kawai, and F. Kano, *Phys. Rev. B* 61, 6574 (2000).
98. C. W. White, J. D. Budai, S. P. Withrow, J. G. Zhu, E. Sonder, R. A. Zuhr, A. Meldrum, D. M. Hembree Jr., D. O. Henderson, and S. Prawer, *Nucl. Instrum. Methods B* 141, 228 (1998).
99. G. Marichy, G. Chassagne, and D. Durand, *Phys. Status Solidi B* 92, 221 (1979).
100. N. E. Christensen and D. L. Novikov, *Phys. Rev. Lett.* 86, 1861 (2001).
101. V. G. Vaks, M. I. Katsnelson, V. G. Koreshkov, A. I. Likhtenstein, O. E. Parfenov, V. F. Skok, V. A. Sukhoparov, A. V. Trefilov, and A. A. Chernyshov, *J. Phys.: Condens. Matter* 1, 5319 (1989).
102. A. van Veen, M. A. van Huis, A. V. Fedorov, H. Schut, C. V. Falub, S. W. H. Eijt, F. Labohm, B. J. Kooi, and J. Th. M. De Hosson, *Mater. Res. Soc. Symp. Proc.* 647, O11.24 (2001).
103. A. V. Fedorov, A. van Veen, and H. Schut, *Mater. Sci. Forum* 363–365, 646 (2001).
104. A. Ueda, R. Mu, Y.-S. Tung, M. Wu, W. E. Collins, D. O. Henderson, C. W. White, R. A. Zuhr, J. D. Budai, A. Meldrum, P. W. Wang, and Xi Li, *Nucl. Instrum. Methods B* 141, 261 (1998).
105. M. J. Puska, P. Lanki, and R. M. Nieminen, *J. Phys.: Condens. Matter* 1, 6081 (1989).
106. P. Asoka-Kumar, K. G. Lynn, and D. O. Welch, *J. Appl. Phys.* 76, 4935 (1994).
107. A. H. Weiss and P. G. Coleman, in "Positron Beams and Their Applications" (P. Coleman, Ed.), p. 129. World Scientific, Singapore, 2000.
108. P. Sferlazzo, S. Berko, and K. F. Canter, *Phys. Rev. B* 35, 5315 (1987).
109. Y. Nagashima, Y. Morinaka, T. Kurihara, Y. Nagai, T. Hyodo, T. Shidara, and K. Nakahara, *Phys. Rev. B* 58, 12676 (1998).
110. For a recent review, see O. K. Andersen, O. Jepsen, and M. Sob, in "Electronic Band Structure and Its Applications" (M. Yussouff, Ed.), p. 1. Springer, Berlin, 1987.
111. H. L. Skriver, "The LMTO Method." Springer, Berlin, 1984.
112. M. Treilleux and G. Chassagne, *J. Phys.* (*Paris*) 40, L-283 (1979); M. Treilleux and G. Chassagne, *J. Phys. Coll.* (*Paris*) 41C, C6/391 (1980).
113. E. Boro'nski and R. M. Nieminen, *Phys. Rev. B* 34, 3820 (1986).
114. B. Barbiellini, M. J. Puska, T. Torsti, and R. M. Nieminen, *Phys. Rev. B* 51, 7341 (1995); B. Barbiellini, M. J. Puska, T. Korhonen, A. Harju, T. Torsti, and R. M. Nieminen, *ibid.* 53, 16201 (1996).
115. M. J. Puska, S. Mäkinen, M. Manninen, and R. M. Nieminen, *Phys. Rev. B* 39, 7666 (1989).
116. D. M. Roessler and D. R. Huffman, in "Handbook of Optical Constants of Solids II" (E. D. Palik, Ed.), p. 919. Academic Press, London, 1991.
117. U. Von Barth and L. Hedin, *J. Phys. C* 5, 1629 (1972).
118. J. Kuriplach, M. i Sob, G. Brauer, W. Anwand, E.-M. Nicht, P. G. Coleman, and N. Wagner, *Phys. Rev. B* 59, 1948 (1999).
119. C. V. Falub, S. W. H. Eijt, A. van Veen, P. E. Mijnarends, and H. Schut, *Mater. Sci. Forum* 363–365, 561 (2001).
120. B. J. Kooi, A. van Veen, J. Th. M. de Hosson, H. Schut, A. V. Fedorov, and F. Labohm, *Appl. Phys. Lett.* 76, 1110 (2000).
121. M. A. van Huis, A. V. Fedorov, A. van Veen, F. Labohm, H. Schut, P. E. Mijnarends, B. J. Kooi, and J. Th. M. De Hosson, *Mater. Sci. Forum* 363–365, 448 (2001).
122. R. Vergara, R. Gonz'alez, J. E. Ramírez, J. E. Muñoz Santiuste, and Y. Chen, *Phys. Rev. B* 58, 5236 (1998).
123. Y. Tanaka, Y. Sakurai, A. T. Stewart, N. Shiotani, P. E. Mijnarends, S. Kaprzyk, and A. Bansil, *Phys. Rev. B* 63, 045120 (2001).
124. The anisotropic part of the 2D-ACAR distribution was determined by subtracting an isotropic distribution that everywhere remains within the measured distribution. The anisotropy is therefore everywhere positive.
125. M. Farjam and H. B. Shore, *Phys. Rev. B* 36, 5089 (1987).
126. M. Forster, J. N. Mundy, and H.-E. Schaefer, in "Positron Annihilation" (L. Dorikens-Vanpraet, M. Dorikens, and D. Segers, Eds.), p. 833. World Scientific, Singapore, 1989.
127. R. Pareja, M. A. Pedrosa, and R. Gonzalez, in "Positron Annihilation" (P. C. Jain, R. M. Singru, and K. P. Gopinathan, Eds.), p. 708. World Scientific, Singapore, 1985.
128. D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.* 45, 566 (1980).
129. J. P. Perdew and A. Zunger, *Phys. Rev. B* 23, 5048 (1981).

130. B. K. Panda, G. Brauer, W. Skorupa, and J. Kuriplach, *Phys. Rev. B* 61, 15848 (2000).
131. P. Hautojärvi and C. Corbel, in [8], p. 491.
132. P. Asoka-Kumar, K. G. Lynn, and D. O. Welch, *J. Appl. Phys.* 76, 4935 (1994).
133. L. J. Romana, J. Rankin, J. R. Brewster, L. A. Boatner, and A. M. Williams, *J. Appl. Phys.* 80, 2781 (1996).
134. J. H. Evans and D. J. Mazey, *J. Phys. F* 15, L1 (1985); *Scripta Metall.* 19, 621 (1985).
135. J. H. Evans and D. J. Mazey, *J. Nucl. Mater.* 138, 176 (1986).
136. Z. Tan, J. I. Budnick, D. M. Pease, and F. Namavar, *Phys. Rev. B* 43, 1987 (1991).
137. M. G. Norton, C. B. Carter, E. L. Fleischer, and J. W. Mayer, *J. Mater. Res.* 7, 3171 (1992).
138. R. C. Birtcher, S. E. Donnelly, M. Song, K. Furuya, K. Mitsuishi, and C. W. Allen, *Phys. Rev. Lett.* 83, 1617 (1999).
139. Y. Nagai, T. Chiba, Z. Tang, T. Akahane, T. Kanai, M. Hasegawa, M. Takenaka, and E. Kuramoto, *Phys. Rev. Lett.* 87, 176402 (2001).
140. M. A. van Huis, A. van Veen, H. Schut, C. V. Falub, S. W. H. Eijt, P. E. Mijnarends, and J. Kuriplach, *Phys. Rev. B* 65, 085416 (2002).
141. C. V. Falub, P. E. Mijnarends, S. W. H. Eijt, M. A. van Huis, A. van Veen, and H. Schut, *Phys. Rev. B* 66, 075426 (2002).
142. M. J. Puska and R. M. Nieminen, *Phys. Rev. B* 46, 1278 (1992).
143. K. O. Jensen, M. Eldrup, N. J. Pedersen, and J. H. Evans, *J. Phys. F* 18, 1703 (1988); K. O. Jensen, M. Eldrup, S. Linderoth, and J. H. Evans, *J. Phys.: Condens. Matter* 2, 2081 (1990).
144. D. T. Britton, P. C. Rice-Evans, and J. H. Evans, *Philos. Mag. A* 55, 347 (1987).
145. Y. Qian, D. Ila, R. L. Zimmerman, D. B. Poker, L. A. Boatner, and D. K. Hensley, *Nucl. Instrum. Methods B* 127/128, 524 (1997).
146. H. Lüth, "Surfaces and Interfaces of Solid Materials," 3rd ed., Sec. 3.1, p. 81. Springer, Berlin, 1995.
147. M. Backhaus-Ricoult, *Acta Mater.* 49, 1747 (2001).
148. G. W. Watson, E. T. Kelsey, N. H. de Leeuw, D. J. Harris, and S. C. Parker, *J. Chem. Soc. Faraday Trans.* 92, 433 (1996).
149. A. T. Macrander, *Phys. Rev. B* 21, 2549 (1980).
150. C. Ronchi, *J. Nucl. Mater.* 96, 314 (1981).
151. M. B. Taylor, C. E. Sims, G. D. Barrera, N. L. Allan, and W. C. Mackrodt, *Phys. Rev. B* 59, 6742 (1999).
152. H. Wengeler and F. Freund, *Ceram. Forum Int. - Ber. Dt. Kerm. Ges.* 58, 31 (1981).
153. A. van Veen, M. A. van Huis, A. V. Fedorov, H. Schut, F. Labohm, B. J. Kooi, and J. Th. M. De Hosson, *Nucl. Instrum. Methods B* 191, 610 (2002).
154. S. J. Zinkle and C. Kinoshita, *J. Nucl. Mater.* 251, 200 (1997).
155. R. N. West, in "Positron Spectroscopy of Solids" (A. Dupasquier and A. P. Mills, Jr., Eds.), p. 75. IOS Press, Amsterdam, 1995.
156. S. W. H. Eijt and C. V. Falub, private communication.
157. M. J. Puska, P. Lanki, and R. M. Nieminen, *J. Phys.: Condens. Matter* 1, 6081 (1989).
158. O. V. Boev, M. J. Puska, and R. M. Nieminen, *Phys. Rev. B* 36, 7786 (1987).
159. B. K. Panda, G. Brauer, W. Skorupa, and J. Kuriplach, *Phys. Rev. B* 61, 15848 (2000).
160. J. Kuriplach, W. Anwand, G. Brauer, and W. Skorupa, *Appl. Surf. Sci.*, in press.
161. A. Nagashima, N. Tejima, Y. Gamou, T. Kawai, and C. Oshima, *Surf. Sci.* 357–358, 307 (1996).
162. N. D. Lang and A. R. Williams, *Phys. Rev. B* 25, 2940 (1982).
163. J. Hulse, J. Kuppers, K. Wandelt, and G. Ertl, *Appl. Surf. Sci.* 6, 453 (1980).
164. K. Jacobi, *Surf. Sci.* 192, 449 (1987).
165. H. Namba and Y. Murata, *J. Phys. Soc. Japan* 53, 1888 (1984).
166. A. P. Mills, Jr. and E. M. Gullikson, *Appl. Phys. Lett.* 49, 1121 (1986).
167. C. C. Griffioen, J. H. Evans, P. C. de Jong, and A. van Veen, *Nucl. Instrum. Methods B* 27, 417 (1987).
168. S. M. Myers, D. M. Bishop, D. M. Follstaedt, H. J. Stein, and W. R. Wampler, *Mater. Res. Soc. Symp. Proc.* 238, 549 (1996).
169. H. Schut, A. van Veen, F. Labohm, A. V. Fedorov, E. A. C. Neeft, and R. J. M. Konings, *Nucl. Instrum. Methods B* 147, 212 (1999).
170. F. W. Clinard, G. F. Hurley, and L. W. Hobbs, *J. Nucl. Mater.* 108–109, 655 (1982).
171. F. Freund, R. Knobel, H. Kathrein, and H. Wengeler, *Nucl. Instrum. Methods B* 1, 223 (1984).
172. S. J. Zinkle and S. Kojima, *J. Nucl. Mater.* 179–181, 395 (1991).
173. W. E. Lee, M. L. Jenkins, and G. P. Pells, *Philos. Mag. A* 51, 639 (1985).
174. Y. Katano, H. Ohno, and H. Katsuta, *J. Nucl. Mater.* 155–157, 366 (1988).
175. P. W. Tasker, *J. Phys. C* 12, 4977 (1979).
176. G. W. Watson, E. T. Kelsey, N. H. de Leeuw, D. J. Harris, and S. C. Parker, *Faraday Trans.* 92, 433 (1996).
177. "Electron Microscopy of Thin Crystals" (P. B. Hirsch, A. Howie, R. B. Nicholson, D. W. Pashley, and M. J. Whelan, Eds.). Butterworths, London, 1965.
178. Y. Qian, D. Ila, R. L. Zimmermann, D. B. Poker, L. A. Boatner, and D. K. Hensley, *Nucl. Instrum. Methods B* 127, 524 (1997).
179. J. H. Evans, *J. Nucl. Mater.* 210, 21 (1994).
180. A. Briggs, *J. Mater. Sci.* 10, 729 (1975).
181. E. A. Kotomin, P. W. M. Jacobs, N. E. Christensen, T. Brudevoll, M. M. Kuklja, and A. I. Popov, *Defect Diffusion Forum* 143–147, 1231 (1997).
182. E. A. Kotomin and A. I. Popov, *Nucl. Instrum. Methods B* 141, 1 (1998).
183. M. Leslie, Program Cascade, Description of Data Sets for Use in Crystal Defect Calculations, Technical report DL/SCI/TM31T, SERC Daresbury Laboratory, 1982.
184. R. W. Grimes, *J. Am. Ceram. Soc.* 77, 378 (1994).
185. R. W. Grimes and C. R. A. Catlow, *J. Am. Ceram. Soc.* 73, 3251 (1990).
186. J. H. Harding, *Rep. Progr. Phys.* 53, 1403 (1990).
187. R. W. Grimes, R. H. Miller, and C. R. A. Catlow, *J. Nucl. Mater.* 172, 123 (1990).
188. H. Schut, A. van Veen, F. Labohm, A. V. Fedorov, E. A. C. Neeft, and R. J. M. Konings, *Nucl. Instrum. Methods B* 147, 212 (1999).
189. A. van Veen, *NATO ASI Ser. B* 278, 41 (1991).
190. S. Godey, E. Ntsoenzok, T. Sauvage, A. van Veen, F. Labohm, M. F. Beaufort, and J. F. Barbot, *Mater. Sci. Eng. B* 73, 54 (2000).
191. C. C. Griffioen, J. H. Evans, P. C. de Jong, and A. Van Veen, *Nucl. Instrum. Methods B* 27, 417 (1987).
192. S. M. Myers, D. M. Bishop, D. M. Follstaedt, H. J. Stein, and W. R. Wampler, *Mater. Res. Soc. Symp. Proc.* 283, 549 (1993).
193. P. F. P. Fichtner, J. R. Kaschny, R. A. Yankov, A. Muecklich, U. Kreissig, and W. Skorupa, *Appl. Phys. Lett.* 70, 732 (1997).
194. D. M. Follstaedt, S. M. Myers, G. A. Petersen, and J. W. Medernach, *J. Electron. Mater.* 25, 151 (1996).
195. J. R. Krasny, P. F. P. Fichtner, A. Muecklich, U. Kreissig, R. A. Yankov, and W. Skorupa, *Nucl. Instrum. Methods B* 136–138, 583 (1998).
196. F. Roqueta, A. Grob, J. J. Grob, R. Jerisan, J. P. Stoquert, and L. Ventura, *Nucl. Instrum. Methods B* 147, 298 (1999).
197. P. F. P. Fichtner, J. R. Kaschny, A. Kling, H. Trinkaus, R. A. Yankov, A. Muecklich, U. Kreissig, W. Skorupa, F. C. Zawislak, L. Amaral, M. F. da Silva, and J. C. Soares, E-MRS, Spring 1998.
198. A. van Wieringen and N. Warmoltz, *Physica* 22, 849 (1956).
199. A. van Veen, in "Fundamental Aspects of Inert Gases in Solids" (S. E. Donnelly and J. H. Evans, Eds.), pp. 41–57. Plenum, New York, 1991.
200. A. van Veen, A. H. Reader, D. J. Gravesteijn, and A. A. van Gorkum, *Thin Solid Films* 241, 206 (1993).
201. R. L. Mills, D. H. Liebenberg, and J. H. Bronson, *Phys. Rev. B* 21, 5137 (1980).
202. H. Trinkaus, *Radiation Effects* 78, 189 (1983).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanorobotics and Nanomanipulation

Wen J. Li

The Chinese University of Hong Kong, Hong Kong SAR, People's Republic of China

Ning Xi, Wai-Keung Fung

Michigan State University, East Lansing, Michigan, USA

Tak Sing Wong

The Chinese University of Hong Kong, Hong Kong SAR, People's Republic of China

CONTENTS

1. INTRODUCTION

Eventually, nanometer-scale devices or systems with dimensions comparable to those of viruses, molecules, or even atoms will become prevalent in all aspects of our lives. For instance, extremely useful applications such as artificial cells or cell repair robots will make it possible to replace or fix damaged cells caused by incurable diseases known today, and to extend human life expectancy. The concept of using engineered protein nanorobots to assemble and manufacture nanodevices or systems was pointed out as early as 1986 by Drexler's landmark book *Engines of Creation* [1] for nanotechnology. He argued that although proteins are not necessarily the best material in terms of mechanical strength or survivability in dry air for building physical nanosystems, it is conceivable that protein-based nanorobots can be used to autonomously build nanosystems out of molecules of much more sustainable materials such as metals and polymers. His concept is analogous to how human hands, although not very strong mechanically nor sustainable in very harsh environments, are able to forge and create much stronger and sustainable machines such as cars and airplanes. In 1992, Drexler [2] actually laid down the foundation for what scientists and engineers refer to today as "nanorobots" by theoretically showing that many different nanorobotic, manipulation, and manufacturing system designs can be realized based on existing technologies such as micromachining, protein engineering, and polymer synthesis.

However, many of Drexler's nanorobots have yet to be realized at the writing of this chapter although it is now generally accepted that one of the exciting prospects of nanotechnology is the possibility of manipulating nanoscale entities such as nanotubes and molecules in a controlled manner to produce novel devices or systems. In this chapter, we will define collectively the developmental efforts that will make this possible: "Nanorobotic and nanomanipulation" technologies, that is, techniques that will eventually allow us to move and manipulate nanoscale (10^{-9} meter) entities such as nanowires/nanotubes (e.g., see [3, 4]), nanocubes (e.g., see [5]), and molecules (e.g., see [6]) or atoms (e.g., see [7]), and assemble them into nanosystems. As an example of nanoscale robotics, it is now possible to fabricate nanoelectromechanical systems which are submicrometer machines that can be actuated through an electrical input signal to produce mechanical movement [8]. These machines can definitely be used to manipulate nanobuilding blocks in the near future.

In the future, "nanorobotics" may be extended to include the construction of robots using nanosize components with overall dimensions at the nanoscale or microscale. However, construction of "robots" at the nanoscale will be very different from constructing conventional robots due to the different dominant physical forces at the nanoscale. (For discussion of size effects on physical forces, see Fearing's survey for microscale forces [9], Sitti's survey for nanoscale forces [10], and Drexler's for a comparison of other physical

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (351–365)

parameters [2]). Nevertheless, simple nanomechanical actuators (e.g., see [11] in 2003) and nanoscale motors (e.g., see [12] in 2000) already have been demonstrated; hence, it is conceivable that nanoscale robots can be constructed soon. As Requicha has pointed out [13], nanorobots are initially likely to be very simple and of limited capabilities. However, a large number of nanorobots can have a significant effect and can tackle complicated tasks. How are such "swarms" of robots to be controlled and programmed? This is an area of work that falls within distributed robotics and so-called artificial life [14].

It is conceivable that once nanoscale construction can be performed by programmable nanorobots, building of nanoscale devices or systems will be done much more effectively at much lower costs. Also, newer designs of these nanodevices or systems can be built and tested more rapidly. Hence, the construction of the first fully controllable or autonomous nanorobot will be a very significant milestone in global research. Examples of groups working toward this goal include Requicha [13] and Cavalcanti [15].

However, most autonomous or controllable nanoassembly and nanomanufacturing techniques are speculative at this time. Even though the potential advantages of nanoscale machinery are obvious and well recognized, practical plans for developing it from existing technology are rare. Why is it currently difficult to build nanoscale machinery? As pointed out by some experts who are trying to commercialize nanotechnology (e.g., see [16]), the main obstacle to develop nanomachines quickly is not a lack of suitable molecular components or methods for connecting them, but simply the fact that there is no existing general technique to bring molecules or nanoentities together and join them into a pattern as specified by a designer. Using existing techniques, when nanoentities are brought together, they tend to assemble in a random fashion or into what their natural affinities would produce (e.g., self-assembly). Hence, there are no "nanohands" to pick up two molecules, hold them, and join them together in a desired manner. This is like the "chicken-or-the-egg" problem; that is, in order to build nanomachines or nanorobots using "nanohands" or tools, one must first have at least a nanomachine to build a "nanohand."

Currently, there are several techniques for building general nanoscale structures, including nanoscale lithography [17], molecular imprinting [18], scanning probes used as "dip pens" [19], and self-assembly [20]. But none of these methods solve the basic problem of assembling preexisting nanostructures or molecular building blocks. They also are incapable of building atomically precise structures, so they may not be useful in solving the "chicken-or-the-egg" problem in building nanomachines.

It is arguable that biotechnology methods will possibly permit the fastest development of molecular machinery since much research effort has already been devoted to protein manipulation and engineering. For instance, self-assembly guided by DNA hybridization [21] is a possible means of solving this problem. In theory, DNA sequences attached to a molecular building block can be redesigned completely independent of the other properties of that building block material; thus novel, useful protein links can be created in this manner.

This chapter will focus on describing only technologies that are mature enough in the sense that they are generally considered by researchers worldwide whenever nanoscale manipulation is required. Excluded from this discussion are the manipulation of nanoentities by chemical means (e.g., self-assembly) or biotechnology methods (e.g., DNA-based molecular building blocks), since they more related to chemical and biological interactions rather than conventional "robotic interactions" (i.e., using controllable physical forces to manipulate matter). So this chapter will present the following topics in detail: (1) manipulation by atomic force microscope (AFM) tips, (2) manipulation by laser trapping force, and (3) manipulation by electrokinetic forces. For a discussion of each technique, the basic principles of operation are first given, followed by a comprehensive survey of past and ongoing activities in using the described method in manipulating nanoentities. Since the manipulation strategy for biological, chemical, or physical nanoentities differ, an attempt to distinguish examples of manipulation work into various categories will be discussed. However, as researchers look into and beyond these nanoscale worlds, the definition of what is biological, chemical, or physical becomes blurred, as matters of all these worlds are made of the elemental entities.

2. SCANNING FORCE MICROSCOPY ROBOTIC NANOMANIPULATION

In the late 20th century, the invention of scanning probe microscopy (SPM) [22] techniques provided a reliable way for humans to visit the nanoscale world, including even the atomic scale world. We now can "see" and also "feel," "touch," and modify nanoscale samples by using recently developed SPM based nanomanipulator systems with the help of the technology of virtual reality (VR), augmented reality, and robotics. This section will focus on the development of nanomanipulation research in the past 20 years.

2.1. Basic Principles

Scanning probe microscopy [22] is a family of microscopy techniques that generate three-dimensional (3D) topography and images of various physical properties (including surface roughness, surface potential, local resistivity, magnetic field, and so on) of samples with resolution close to atomic scale. In general, a sharp tip is rastered across the surface of samples under investigation while tip–sample interactions are monitored. Different forms of microscopy are characterized by the kind of tip–sample interactions involved, that is, scanning tunneling microscopy (STM), AFM, lateral force microscopy, electric force microscopy, near-field scanning optical microscopy, etc. [22]. The most commonly employed forms of SPM are STM and AFM and most of the commercially available SPM systems are designed to support various microscopy forms by installing suitable tips and accessories. The SPM system was originally designed for imaging purposes only, for studying surface profile of various physical properties of samples with nanometer scale resolution. Since the invention of STM and AFM systems, researchers'

attention has been increasingly attracted to modify sample surfaces and manipulate atoms, molecules, and biological entities. Solving technical problems involved in the use of STM and AFM as nanomanipulators became the major research direction in nanorobotics.

Scanning tunneling microscopy was developed by Binning et al. at IBM Research Laboratory in Zurich, Switzerland in 1982 [22]. In general, STM consists of a metallic sharp probe (usually made from tungsten, nickel, or gold) mounted on a small cantilever (with a length of about 100–200 μm), a piezoelectric scanner that provides precise motions in X-, Y-, and Z-directions, and electronics to control the motion of the piezoelectric scanner. The cantilever with a sharp probe is mounted on the 3D piezoelectric scanner and is scanned across on the sample surface in a raster manner. During the raster scan, small tunneling current between the tip and sample surface is monitored so that the distance between the tip and sample surface can be measured. The tunneling current increases exponentially as the distance between the STM tip and sample surface decreases and thus resolution less than 1 Å can be detected for the measurement. By keeping the tunneling current constant by a feedback loop during scanning, the Z-position of the tip is continually adjusted by the piezoelectric scanner. The Z-displacements of the piezoelectric scanner during raster scan are recorded and a 3D topographic map of the sample surface can be generated. The 3D surface topography obtained from STM actually reveals the structure of electron clouds of individual atoms. One of the major limitations of STM is that only conducting materials (i.e., metals and semiconductors) can be imaged using STM. Biological molecules and entities and other nonconducting materials cannot be imaged using STM.

In order to remedy the aforementioned limitations of STM, atomic force microscopy was developed by Binning et al. as a collaboration of IBM Research Lab, Zurich and Stanford University in 1986 [23]. The basic setup of AFM is similar to that of STM, but with different tips, which are typically MEMS*-fabricated with silicon (Si) or silicon nitride (Si_3N_4). During raster scanning of the sample surface, the Z-displacement of the piezoelectric scanner is adjusted by maintaining constant deflection of the cantilever and thus the atomic force exerted on the tip by a feedback loop. The cantilever deflection is detected optically by collecting reflected laser off the cantilever (which is coated by gold) using a position sensitive detector (PSD) and a differential amplifier. The PSD usually consists of two or four photodiodes to detect the minute changes in the light path of the laser which reveal the deflection and twist of the cantilever, as shown in Figure 1. Contrary to STM which can only image conducting materials, AFM can image all kinds of materials, including nonconducting materials, and can work under fluid. This allows biological entities, like cells, proteins, DNA, organelles, bacteria, viruses, etc., and any surfaces that involve a solid–liquid interface, like geologic samples, corrosive surfaces, etc., to be imaged with nanometer scale resolution. The diversity of samples that are suitable to be imaged using AFM promotes the popularity of AFM in nanotechnology. There are three operating

* MEMS stands for MicroElectromechanical System

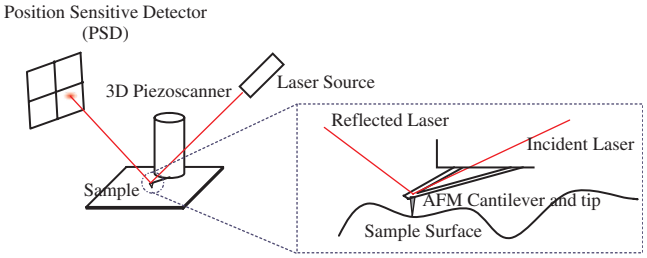


Figure 1. Schematics of an atomic force microscope. (Components are not drawn in scale.)

modes of common AFM systems, including contact mode, noncontact mode, and TappingMode™. In contact mode, the AFM tip has physical contact with the sample surface and thus contact mode can give the highest image resolution and scan rate among the three AFM operating modes. However, fragile samples, especially biological samples, may be damaged by the AFM tip under contact mode. On the other hand, the AFM tip does not have physical contact with samples under the noncontact mode and thus no damage is imposed on fragile samples. The image resolution obtained and scan rate under noncontact mode are, however, the lowest among the three modes. TappingMode™ scanning imposes minimal damage on fragile samples while maintaining high resolution of AFM images. In TappingMode™, the cantilever is oscillated at or near its resonant frequency during scanning across a sample surface and thus the cantilever taps the sample. The oscillation amplitude can range from 20 to 100 nm and the tapping force is less than 200 pN. The Z-displacement of the piezoelectric scanner is adjusted by maintaining a constant root-mean-square (rms) value of oscillation signal captured by the PSD. Since AFM can image samples of all materials and under fluids, it has become a good candidate as a nanomanipulator instead of STM recently. Both contact mode and TappingMode™ AFM can be employed to manipulate nanosized objects.

Various types of manipulation tasks are performed using AFM and STM systems. Mechanical tasks, including pushing/pulling, orienting, bending, twisting, rolling, sliding, picking-and-placing, assembling, cutting, indenting, lithography, and so on, can be achieved using STM and AFM. SPM systems can also be modified to the so-called "specific" manipulators, which only respond to particular types of molecules by coating the SPM tip with layers of catalysts, enzymes, antigens, and any chemical/biochemical molecules that only are reactive with or respond to the nano-objects of the targeted types and properties. Antigen–antibody interaction is one of the most commonly adopted mechanisms for making "specific" end-effectors and performing selective *in-situ* manipulation. Detailed discussion of the recent research results will be given in the following sections.

2.2. Early Attempts at Atomic and Molecular Manipulation

When people first employed STM or AFM to image sample surfaces at atomic scale resolution, the capability of modifying and manipulating sample surfaces of the interaction of the SPM tip (electric field and force generated by the STM and the contact force exerted on the sample surface by the AFM) was obviously observed. Many

researchers have attempted to control the surface modification or manipulation using SPM in a precise and selective fashion since the 1980s. The early attempts of surface modification experiments involved unstructured deposition or removal of atoms or clusters of atoms from sample surfaces using STM although the results were inconsistent and in an uncontrolled manner. Some of the pioneer works are listed. Becker et al., of AT&T's Bell Laboratories, created nanometer sized features on a germanium (Ge) surface by applying short and high voltage pulses using a STM tip [24]. The average size of the features is about 0.8 nm in diameter and 0.1–0.2 nm in height. Germanium atoms were adhered to the STM tip and transferred to the sample surface during each pulse while scanning. Mamin et al., from IBM's Almanden Research Center, emitted gold atoms from a STM tip to form small gold clusters on gold, glass, and platinum surfaces under vacuum by applying high voltage to the STM tip [25, 26]. On the contrary, Clery, from Hitachi Central Research Laboratory, was able to remove sulfur (S) atoms from a molybdenum disulfide (MoS_2) crystal using voltage pulses applied a STM tip so that letters and characters patterns with heights of 1.5 nm were etched out from the surface [27]. Other researchers have also made progress in the techniques of surface modification, as reported in [28–30].

As more and more attempts in surface modification in the nanometer scale have been conducted, researchers from various disciplines gained much experience in manipulating atoms and molecules and aimed to perform precise position control and manipulation of atoms and molecules that would be required for assembly of nanodevices. One of the famous examples was demonstrated by Eigler and Schweizer at IBM Almaden Laboratory in 1990 [7]. They positioned 35 Xenon (Xe) atoms to spell out the "IBM" pattern on a nickel (111) surface using a tungsten STM tip under ultrahigh vacuum and low temperature (about 4 K). Their work was the first that demonstrated the ability to manipulate atoms to form structured patterns. They also achieved similar results of manipulation of different atoms (e.g., iron) on different surfaces (e.g., copper) into various patterns (e.g., circle) with other collaborators in following years [31, 32]. Moreover, Durjardin et al. dissociated and examined individual decaboron tetradecahydride molecules ($B_{10}H_{14}$), that were adsorbed on a silicon (Si) surface, into fragments with low energy electron pulses using STM [33]. This technique can be further exploited in semiconductor manufacturing for nanometer scale site-specific doping of boron (B) in silicon (Si). Mo et al. also reported that antimony (Sb) dimers could be rotated reversibly between two stable states on a silicon (001) surface under room temperature by applying low energy voltage pulses to a STM tip [34].

In summary, the aforementioned examples of manipulation of atoms and molecules were conducted manually. Each manipulation step in these early attempts of surface alteration in the nanometer scale was not easily achieved and repeated iteratively until the desired goals were achieved. The experimental conditions for manipulation of atoms and molecules were usually demanding so that the manipulation tasks were performed by STM with application of high energy voltage pulses under ultrahigh vacuum and low temperature. Some of the experiments were also conducted on a trial-and-error basis. This made early atomic manipulation a low yield, time-consuming, tedious, and highly restricted task, with a lack of precise positional control. The experimental results were also difficult to reproduce. The development of technology of nanomanipulation has then marched to the direction of employing AFM as a nanomanipulator so that the sample type for manipulation can be extended to nonconducting and fragile biological entities.

2.3. Recent Developments in SPM

Inspired by the preliminary results of atomic manipulation obtained in late 1980s and early 1990s, more and more research groups from various disciplines (including physics, chemistry, biotechnology, computer science, robotics, etc.) in the world have joined the exploration of the nanoworld since the mid 1990s. The manipulated objects were not limited to conducting materials and the manipulation environments were not limited to ambient. The application of nanomanipulation has been extended to nonconducting materials and even objects under fluids, like polymers, living cells, DNA, proteins, etc. This led to the trend of using AFM as the major nanomanipulation tool in recent years. In order to assist operators to have better and finer control on the AFM tip (the end-effector), various types of sensor information are fed back and displayed to operators during manipulation and operators can issue motion command to the AFM tip and receive sensor feedback (especially force information) by using haptic devices. The technology of VR has also been introduced into nanomanipulation systems so that operators can have a better understanding of the environment for nanomanipulation. In additions, some of the developed AFM based nanomanipulation systems can also be teleoperated via the Internet. Some of the representative nanomanipulation technologies developed by research groups around the world are listed. One of the pioneering nanomanipulator systems was a home-built AFM developed under the telenanorobotics project of the University of Tokyo [35]. The system was also equipped with a home-built 1 DOF haptic device for normal force feedback from the AFM tip and a VR display as visual feedback. The NanoManipulator™system, integrated by the collaboration of the Departments of Computer Science and Physics of the University of North Carolina at Chapel Hill, consists of an AFM, a haptic device, and a 3D visualization system. Operators can control the real-time 2D motion of an AFM tip and receive surface topography information as force feedback via the haptic device while having synthetic 3D visual feedback of the operating environment [36]. This integrated system has been commercialized in the market. In order to have a better picture of the environment during nanomanipulation, an AFM can be built inside the vacuum chamber of a scanning electron microscope (SEM). Examples include the systems developed at Fukuda and Vachkov's laboratory at Nagoya University [37], Washington University at St. Louis [38], and the Division of Microrobotics and Control Engineering at the University of Oldenburg [39]. Some research groups also modify the end-effector (i.e., the AFM tip) of AFM based nanomanipulators by, for example, integrating micropiezoresistive force sensors [39], using capacitative ultrasonic actuators [40] on AFM cantilevers, and attaching carbon nanotubes to an AFM tip

as nanoscissors [41]. An interferometric system was also installed to a SPM system for high precision position measurement for real-time tip position control [42]. In addition, the nanomanipulator system developed at the Robotics and Automation Laboratory of Michigan State University allowed operators to control 3D motion of the AFM tip and receive 3D force feedback via the Internet in real time [43]. The nanomanipulation system developed by Friedt et al. can also be remotely controlled [44].

2.4. Survey on SPM Nanomanipulation Research

In general, nanomanipulation tasks are divided into two categories, namely 2D and 3D nanomanipulation, respectively. For 2D nanomanipulation, the SPM tip only moves along the substrate surface so that the manipulated entities only translate/slide along and rotate on the operating surface and never move off the surface. On the other hand, the position of the SPM tip can be controlled to move freely in 3D space for 3D nanomanipulation so that the 3D position and even orientation of the manipulated entities can be controlled.

2.4.1. 2D Nanomanipulation

Kim and Lieber demonstrated to the rearrangement of thin oxide structures on the underlying surface by increasing the applied load during scanning with AFM [45]. The sled-type motion of C_{60} islands has been studied by Luthi et al. using AFM [46]. By using AFM, Schaefer et al. deliberatively moved gold (Au) clusters on a smooth surface [47]. Junno et al. demonstrated the formation of arbitrary nanostructures by controlled pushing of individual GaAs particles (with diameter of about 30 nm) under ambient conditions [48]. Requicha et al. of the University of Southern California also demonstrated the construction of arbitrary patterns of gold nanoparticles using AFM under ambient conditions [49] and in liquid environments [50]. The same group also reported translational and rotational manipulation of gold nanorods on a SiO_2 surface [51, 52]. Hertel et al. demonstrated that they could straighten, bend, translate, reorient, and cut carbon nanotubes on glass substrate using contact AFM [53]. They fabricated a complex pattern (e.g., the Greek letter Θ) from a carbon nanotube bundle of 2.5 μm length. They also measured binding energy between carbon nanotubes and substrate surfaces. Postma et al. measured electronic transport on carbon nanotube junctions, buckles, and crossings that were fabricated in a controlled manner from straight and undeformed carbon nanotubes by manipulation using AFM [54]. Falvo et al. demonstrated bending of straight nanotubes into various configurations and measured resistance of multiwalled carbon nanotubes on highly oriented pyrolytic graphite with changing orientation using their AFM based nanomanipulator system with VR capability [55, 56]. Tombler et al. experimentally investigated the reversible metal–semiconducting transition property of carbon nanotubes by pressing on suspended nanotubes using an AFM tip [57]. In these experiments, the samples were first imaged in noncontact mode to minimize the lateral force acting on the samples and then manipulation was carried out with the normal force feedback switched off. The aforementioned manipulation experiments fell in the 2D manipulation category. One problem of this method is that the normal force cannot be controlled during manipulation, as the force feedback in the Z-direction was switched off. This might result in either breaking cantilevers due to large normal force or insufficient force to keep the tip in contact with the surface [36]. Another problem of this manipulation scheme was that it could only manipulate two-dimensional nanostructures on a smooth and flat substrate surface. The surface tilt must also be carefully removed before manipulation. The reported manipulation techniques also could not be extended for the development of nanoassembly techniques, like picking-and-placing nanoparts, mating of the nanoparts, and so on. One solution for the aforementioned problems is to use techniques developed for 3D nanomanipulation so that the 3D position and forces of the tip can be controlled.

2.4.2. 3D Nanomanipulation

There are also ongoing research projects focusing on 3D nanomanipulation. A promising method for 3D nanomanipulation is to build the AFM inside the vacuum capsule of a SEM. Piezoelectric vacuum manipulators constructed inside the SEM have the ability to manipulate objects along the X-, Y-, and Z-directions using the AFM tip as the end-effectors. Several experiments of manipulation of carbon nanotubes were performed using this kind of setup [37, 38, 58]. The obvious advantage of this method is that multiple end-effectors can be built inside the SEM to achieve more degrees of freedom. The manipulation can be performed between the end-effectors without the need for a substrate. The operation can also be monitored in real time. However, the measurement of the deflection of the cantilever is based on images from the SEM, whose resolution is not enough. Thus the manipulation accuracy of this method is also not comparable to that of the AFM. Since the samples are placed in vacuum and exposed to electron beam with high energy, this manipulator cannot be used to manipulate biological samples. The expense of a SEM, ultrahigh vacuum conditions, and space limitations inside the SEM vacuum capsule also impede the wide application of this method.

Different manipulation techniques have also been attempted for realization of 3D nanomanipulation. Hansen et al. proposed in [59] such that the nanoparticles were imaged in TappingMode™and pushed over a surface using contact mode, while the normal force feedback was switched on. This method can be considered as the beginning of the 3D nanomanipulation in the sense that the cantilever tip follows the topography of the surface using the internal feedback control. However, this method takes the risk of hazardous action of the feedback mechanism when switching on and off the vibration of the cantilever, changing the gains of the feedback loop, and changing other parameters such as set point and tip velocity, while the tip is touching the sample surface. Recently, some researchers have combined AFM with VR interface and haptic devices to facilitate nanomanipulation. By introducing a virtual environment of the samples, nanomanipulation using the AFM becomes much easier [36, 55]. Besides, 3D synthetic visual feedback to users, a 1 DOF haptic device had also been constructed in [60] for haptic sensing. In the manipulation mode, during approach and contact of the surface or the

object, or manipulating the object, the operator controls the X–Y–Z motion of the cantilever while feeling the normal tip–sample interaction force. However, because only the normal force is "feelable" during manipulation, this method may also take the risk of breaking the cantilever and damaging the tip when large lateral forces occur. In [36], the AFM tip was also virtually connected to a Phantom™ stylus. In imaging mode, the topography data coming from the AFM are sent to the Phantom™ controller, which reconstructs a "feelable" surface for the user. Then the user can feel the topography of the sample. In manipulation mode, the stylus can be used to move the tip over the surface while keeping the inside normal force feedback on. However, in this approach, the operator does not directly feel forces, which are proportional to the actual forces acting on the cantilever. These two methods can also be considered as the 3D nanomanipulation since the cantilever tip can be controlled either by the internal force feedback control loop or by the operator through the haptic devices to follow the topography of the surface. In addition, Li et al. introduced the concept of an AFM frame to avoid the problem caused by the "bow" effect of the piezoscanner of the AFM system in 3D manipulation [43]. Topography represented in AFM frame is always convoluted with a convex bow. A flat surface in the Cartesian frame will become a concave bow in an AFM frame as shown in Figure 2. The curvature of the bow cannot be negligible in large scan size as the uncorrected bow can be 60 nm for 90 μm $\times$ 90 μm scan size. In order to remedy the bow effect in AFM based 3D nanomanipulation, a 3D trajectory planning algorithm was proposed by Li et al. in [43]. Before 3D trajectory planning for nanolithography, the surface of the nanofixture substrate is first scanned. The 3D topography of the surface is thus represented by a matrix that records the Z-coordinate of sampled points on the sample surface. This matrix thus depicts the relation between the conventional Cartesian frame and the AFM frame with respect to the surface for manipulation. Any 3D trajectory of the AFM tip with respect to the Cartesian frame can then easily be transformed to a representation with respect to the AFM frame by a simple look-up table technique. Three-dimensional nanolithography experiments on tilt surfaces have been conducted to show the effectiveness of the approach. Figure 3 shows two scribed "CUHK" patterns on a tilt surface using an AFM with automatic surface following capability turned off (i.e., the normal force feedback was turn off). The upper pattern was scribed without 3D trajectory planning. It can be seen that the scribed depth of the four letters are different due to the surface tilt. Moreover, the left letter "C," which is scribed on the higher side, was scribed deeper than the right letter "K," which is on the lower side. On the other hand, the lower pattern was scribed with 3D trajectory planning. The depths of the scribed letters are similar in the pattern.

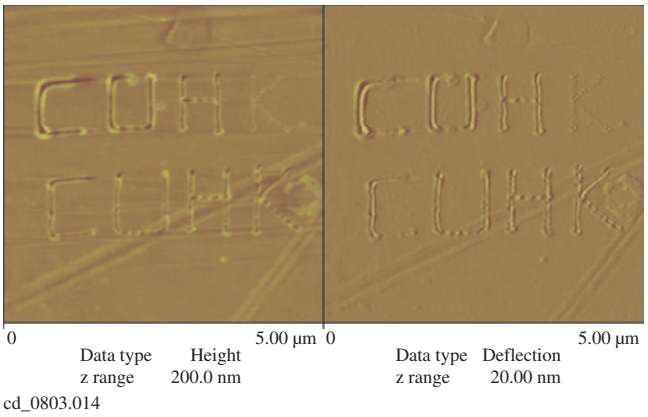


Figure 3. Nanolithography on a tilt surface. Upper: Without 3D trajectory planning, Lower: With 3D trajectory planning. Left: Height signal, Right: Deflection signal.

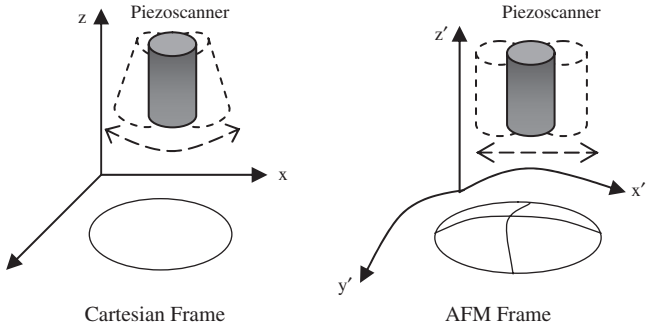


Figure 2. Cartesian frame vs AFM frame. A flat surface with respect to the Cartesian frame is a concave surface with respect to the AFM frame.

2.5. Biological Nanomanipulation

Recently, the development of the technology of nanomanipulation has attracted much attention from the areas of genomics, proteomics, cell biology, and medicine for an exciting future in the investigation of the structures and functioning mechanisms of living things and their interactions at the molecular level [61, 62]. Gutholdm et al. manipulated and studied the DNA molecules and viruses using a nanomanipulator system [63]. Biologists employed an AFM tip as a microdissection tool to reveal minute structures such as organelles and even chromosomes of cells and bacteria. Hoh et al. revealed the hexagonal array of hepatic gap junction hemichannels by applying force on an AFM tip to the top membrane of the gap junction for dissection [64]. Hansma et al. employed an AFM tip as a dissection tool to reveal DNA of plasmid [65]. Stark et al. demonstrated the dissection of human chromosomes in ambient condition and in buffer by AFM [66, 67]. They reported that the dissection results were improved with z-modulation because the lateral force during cutting was reduced with z-modulation of the AFM tip. Precise cutting with a loading force larger than 17 μN with a full width at maximum depth of 90 nm can be achieved with z-modulation. This method could also be employed for chromosomal material extraction because the chromosomal material adhered to the AFM tip after dissection. On the other hand, researchers also "functionalized" the AFM tip to form selective grippers, which only function for specific entities/molecules. This concept was first introduced by Drexler and Foster in 1990 [68]. Specially designed protein molecules and antibodies can be adhered to AFM tip to serve as a gripper for holding or manipulating

molecules in a selective manner. Similarly, site specific reactions can be activated by mounting suitable catalyzes and agents on an AFM tip. Florin et al. measured the binding affinity between receptor avidin molecules and biotin ligand molecules by binding receptor avidin to an AFM tip [69]. Fruitful results are also obtained in antibody recognition and localization [70–72]. The basic idea is to attach suitable antigen molecules on the AFM tip and measure the binding force detected by the deflection of the AFM cantilever during imaging. By analyzing the obtained binding force profile, the type and the location of the antibodies on the sample surface can then identified.

2.6. Summary

In conclusion, the development of SPM based nanomanipulation is still in a preliminary stage. Nanometer to micrometer sized entities, for instance gold nanoparticles, carbon nanotubes, protein molecules, DNA, bacteria and viruses, etc., can be pushed/pulled, slided, rotated, reoriented, picked-and-placed, and dissected using SPM under ambient condition or in liquids manually with the assistance of various kinds of input and visualization devices for sensor feedback. During nanomanipulation, operators can receive real-time force feedback and issue motion commands to the AFM tip via a haptic device and have VR based synthetic visual feedback. Some of the nanomanipulation systems can also be operated remotely via the Internet. However, our current achievement of nanomanipulation still cannot allow us to have mass production of functional nanodevices and nanorobots due to the lack of:

1. a method of precise control (both in 3D position and force) of the SPM tip, in terms of sensor design, control algorithms, and actuation;
2. designs of specific end-effectors for particular tasks;
3. complete and reliable models for different manipulation tasks, such as the nanotribological phenomenon of the interface in pushing objects [73];
4. algorithms and techniques for automated nanoassembly.

3. LASER BASED NANOMANIPULATION

Laser based manipulation of objects refers to techniques of holding or trapping of entities by optical instead of mechanical means. The entities can then be moved or manipulated by controlling the motion of the focus of the laser source using a lens and mirror system. The optical tweezers technique is most commonly employed for laser based nanomanipulation. Optical tweezers provide noncontact manipulation for nanometer to micrometer sized neutral objects in any condition, including ambient and liquid environments. The applied force on the manipulated objects can be as small as pico-Newton range. This makes optical tweezers important for biological and chemical studies of fragile samples. Ashkin, at Bell Laboratory, first demonstrated how to optically accelerate transparent latex spheres, with diameter of 1 μm, suspended in water with a mildly focused 1 W argon laser [74]. The motion of the latex spheres was observed to be guided by the motion of the beam. In 1986, Ashkin et al. developed 3D cooling of dielectric particles using a single laser source [75], which formed a strong foundation for the development of optical tweezers. The technique of optical trapping plays a significant role in the fundamental research of various areas so that S. Chu was awarded the 1997 Nobel Prize in Physics for his contributions in atom cooling and trapping using lasers.

3.1. Basic Principles

Optical tweezers are defined as 3D traps to trap or manipulate micrometer and nanometer sized neutral charged objects by using a focused laser. The source of the trapping ability of optical tweezers comes from the radiation pressure of light. The laser light employed in the optical tweezers carries momentum $p = \frac{h\nu}{\lambda}$, where h is Planck's constant, ν is the frequency of the laser, and λ is the wavelength of the laser. When the laser light hits an object, the laser is either reflected from or refracted (transmitted) through the object. The momentum carried by the transmitted laser light is different from the incident laser light. The change of momentum conveyed by the laser constitutes the radiation pressure. By the law of conservation of momentum, there must be an equal but opposite momentum change in the object. Forces are generated due to the change of momentum. The object can then be trapped when all external forces (for example, gravity and electric force) exerted on the object are balanced by the forces resulting from the change of momentum. Two forces come from the radiation pressure, namely the scattering force F_{scatt} and gradient force F_{grad}, respectively.

Depending on the size of the trapping objects, optical trapping is categorized into two regimes, namely the Rayleigh regime when the wavelength of the light is comparable to object size d or $d \approx \lambda$ and the Mie regime when the wavelength of the light is very small compared to object size, $d \gg \lambda$ [75]. In the analysis of the scattering and gradient forces in optical trapping, the dipole model is employed for objects in the Rayleigh regime which considers the wave properties of light. Dipoles are induced in the objects from the laser. The scattering force F_{scatt} comes from reflected or adsorbed light by the objects and is proportional to the light intensity and in the direction toward the laser propagation. The laser also polarizes the objects and the gradient force F_{grad} is exerted on the objects by the electromagnetic field gradient. The gradient force F_{grad} is proportional to the gradient intensity and points in the intensity gradient direction. The resultant force exerted on the objects then pushes the objects toward the laser focus. On the other hand, the ray-optics model can be employed in analyzing scattering and gradient forces of optical trapping in the Mie regime [75].

Figure 4 depicts the geometrical view of the trapping forces generated when laser light hits a dielectric sphere. The trapping force arises from the refraction of laser light, which generates change of momentum from photons of the laser light. The resultant exerted force, F, on the object brings the object toward the focus, f, of the incident laser light, as shown in Figure 4. In other words, if the center of the object is lower than the laser light focus, the object is pushed upward as shown in Figure 4(a), or vice versa as shown in Figure 4(b). Ashkin derived the mathematical expressions of the involved scattering and gradient forces

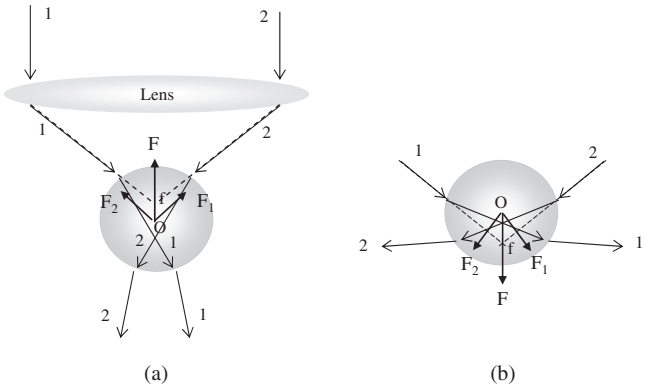


Figure 4. Trapping forces in optical trapping.

for optical trapping in the Mie regime, which depend on the power of incident laser, angles of incidence and refraction, refractive index of the suspending medium, fraction of reflected light intensity from the surface, and fraction of transmitted light intensity in the object [76]. The resultant exerted force, F, on the object is then given as $F = F_{\text{scatt}} + F_{\text{grad}} = Qn_mP/c$, where Q is the dimensionless quality factor, n_m is the refractive index of the suspending medium, c is the speed of light, and P is the incident power of the laser [76]. The quality factor Q depends on the relative refractive index of the object and the suspending medium.

Typical optical tweezers are usually installed on standard optical microscopes and the laser light can be focused by microscope objective. The objects for manipulation are placed on the microscope stage. Typical laser sources are Nd:YAG laser and argon laser with power range from 25 to 500 mW for manipulating biological entities. Typical wavelength of the laser lies around the infrared spectrum (i.e., 700 and 1060 nm). Manipulation of objects, such as moving objects, can be achieved by manipulating the position of the focus of the laser beam using complicated lens and mirror systems.

The primary application of optical tweezers is for holding and manipulating biological entities [77]. With the help of laser scalpel for microdissection, covered structures, like cell organelles, can also be manipulated optically. Ashkin and Dziedzic were the first that demonstrated trapping of living biological entities, including *Escherichia coli* bacteria and viruses in 1987 [78]. In this pioneer work, they showed that laser trapping was harmless to living biological entities with laser wavelength at 1060 nm up to several hours. Reproduction of *Escherichia coli* bacteria and yeast cells within optical traps was also reported. The optical trap generated was strong enough to translate the trapped *Escherichia coli* bacteria with speed up to 500 m s^{-1} [79]. Later research also showed that optical trapping did not damage pigmented red blood cells, green plant cells, algae, and protozoa. Manipulation of organelles and particles with plant cells and protozoa was achieved without damaging the cell membranes and only deformation of the cell shape was observed during manipulation. Aufderheide et al. also achieved similar results by directed positioning of micronuclei and other organelles to any location inside Paramecium tetraurelia bacteria without damaging the bacteria [80].

Optical tweezers were also applied in human reproduction and assisted fertilization research. Colon et al. [81] and Tadir et al. [82] both demonstrated 3D manipulation of living human sperm cells. Microsurgery was conducted for drilling channels in the zona pellucida of oocytes using optical tweezers and a laser scalpel in order to facilitate sperm penetration [83]. Conia and Voelkel also succeeded in manipulation of human gametes with optical trapping [84]. Schütze et al. succeeded in direct insertion of sperms into channels to simulate fertilization [85, 86]. Tadir et al. demonstrated how to modify preimplanted embryo and gametes with optical manipulation [87]. Optical tweezers were also found to be useful in dissection in the micro- or even nanoscale level. Blanchet et al. developed a new zona drilling technique for manipulation of internal structures of mouse embryo [88]. Weber et al. have several achievements in manipulation of cellular organelles and genomes [89]. In particular, they extracted DNA from chloroplasts from a plant cell using a laser microbeam [90] and a dissected part of a human chromosome out for cloning [91]. Liang et al. also performed intracellular surgery on chromosomes in PTK-2 cells using a laser scalpel, and optical tweezers were employed to hold the dissected distal chromosome fragments together throughout anaphase [92]. The same group also dissected and manipulated chromosomes in living newt lung epithelial cells [93]. In their experiments, they demonstrated that the dissected chromosome fragments could be easily pulled or rotated by optical tweezers when positioned at the periphery of the mitotic spindle.

Optical tweezers were applied in characterization of the structure and behaviors of various biological organelles, molecules, and mechanisms. For example, Bayoudh et al. characterized the chloroplast arrangement, shape, and consistency in cells of living leaf tissue [94]. Two optical tweezers were applied to provide two-point contacts to a single chloroplast in suspension. A microstirrer, which was constructed by the third optical tweezers holding a birefringent particle, induced tumbling motion to the suspended chloroplast and its various properties were investigated. Sakata-Sogawa et al. directly stretched looped DNA molecules using two optical tweezers and measured their molecular length [95]. The DNA looping is mainly due to the interaction among DNA binding proteins distributed along the molecular structure of the DNA molecules. The stretched tension force of the DNA molecules was about 30 pN and the measured rise per residue was computed as 3.31 ± 0.05 Å. By using two optical tweezers, Bronkhorst et al. brought two red blood cells together and then pulled them apart for the investigation of the interaction during aggregation (bringing together) and disaggregation (pulling apart) [96]. Viegal et al., in 1998, measured the stiffness of rabbit skeletal actomyosin cross-bridges using two independent optical tweezers [97]. The two ends of the actin filament were first attached to two latex beads, which were individually trapped by two optical tweezers. A third bead, which was coated with myosin, was then placed and fixed near the actin filament in order to induce a working stroke of 5.5 nm to the system. The stiffness of the cross-bridge system was then computed based on the displacements of the latex beads, which were measured by interferometry. In addition, Bennink et al. characterized the interaction mechanism of

DNA with RecA protein and YOYO-1 dye molecules using single optical tweezers [98]. Both the optical tweezers and a micropipette first captured polystyrene beads individually. Then, a DNA molecule was attached to the beads in stranded configuration by introduction of DNA buffer solution to the setup. The micropipette then moved the bead to introduce deflection of the laser beam. The displacement of the laser trapped bead could be measured, which in turn determined the force exerted to the system by moving the bead by micropipette. Elasticity and other structural parameters of in the DNA-RecA protein interaction system could be computed.

Another type of application of optical tweezers in biological applications is separation and sorting of interested biological entities out of a mixture of similar entities. Early attempts included that by Buican et al. who developed an automated technique to identify and pick up the interested cell [99] using optical trapping. Separation of bacteria in unidentified water and soil samples was demonstrated by Mitchell et al. [100] and Huber et al. [101] by using optical tweezers.

3.2. Manipulation of Chemical Entities

The primary research in manipulation of chemical entities focuses on cooling or trapping of atoms using a focused laser, i.e., confining a dense cloud of atoms in a small volume [102–104]. Furukawa and Yamaguchi reported that gold particles (with diameters in the range of 0.5–3.0 μm) suspended in water could be trapped using single optical tweezers [105]. Svoboda and Block also reported similar results [106]. Their work gave the evidence that metallic particles could also be trapped optically. Friese et al. demonstrated that micrometer-sized calcium carbonate ($CaCO_3$) particles could be either aligned along the polarization plane of the incident laser or rotated with constant speed [107]. A rotation frequency of 357 Hz could be obtained with a 1 μm thick particle and a laser beam with a power of 300 mW. Recently, Paterson et al. demonstrated controlled rotation of microscopic silica spheres, glass rods, and chromosomes that were trapped by optical tweezers with a spiral interference pattern from two laser sources [108]. The rotation speed could be higher than 5 Hz. The interferometric pattern was generated by interfering an annular shaped laser beam with a reference beam. The rotation of objects was due to the difference in optical path length of the two laser beams. The same group then constructed 3D stacked structures, which were made of microspheres with diameters of 53 μm, by the introduction of an interferometric pattern between two annular laser beams [109]. The constructed structures can be translated in 3D with a speed of 200 $\mu m\,s^{-1}$ using 200 mW laser beams and can be rotated continuously by introducing frequency difference between the two laser beams. The structures of the constructed microspheres stacks could be controlled, including asymmetric structures, by the geometry of the interference patterns of the two laser beams. Moreover, Diddams et al. developed an all-optical atomic clock by optically trapping a single $^{199}Hg^+$ ion [110]. The laser source employed in this clock was a mode-locked femtosecond (10^{-15} s) laser with output pulses at 1 GHz rate and was phase-coherently locked to the frequency of the laser light. The ultimate fractional frequency uncertainly was expected to be approaching 10^{-18}, which was much smaller than that (10^{-15}) of the world's most accurate caesium (Cs) based clock. Optical trapping of ions can be applied in realization of quantum computing systems. Duan et al. proposed an implementation scheme for quantum computation based on the Abelian and non-Abelian geometric operations by manipulating a set of trapped ions using laser [111]. In 2001, Kielpinski et al. demonstrated an implementation of memory of one qubit (the basic unit memory for quantum computing) [112]. The information of one qubit was encoded by manipulation of a pair of trapped $^9Be^+$ ions.

3.3. Summary

In conclusion, the development of micro- and nanoscale manipulation technology using laser trapping is still in the preliminary stage although extensive results in the biological areas have already been obtained. Applications of optical trapping in biological areas have been developing rapidly since the 1990s and are expected to grow in the coming decades. However, few works focus on the assembly of microdevices and nanodevices using optical tweezers. The lack of precise position control algorithms for manipulating optically trapped objects and the small workspace of optical tweezers hinder the development of an automated assembly plan using optical trapping.

4. ELECTROKINETIC BASED NANOMANIPULATION

Electrokinetic based manipulation has been considered as one of the most versatile methods in microscale and nanoscale manipulation. Research in using electrokinetic forces to manipulate microscale and nanoscale entities has been conducted actively in the past decade due to the advancement of micro/nanofabrication technology to fabricate micro/nanoscale electrodes for the generation of ultraintense electric field ($>10^6$ V/m) by small input electrical excitation. Electrokinetic forces (EK forces) generally refer to the Columbic forces imparted on particles when external electric field is applied. Precise classification of EK forces is defined according to the mechanisms by which the forces are generated. For instance, *dielectrophoretic* force is defined as the force that is generated when a neutral particle is immersed in a nonuniform ac electric field region while *electrophoretic* force is defined as the force that is generated when a charged particle is immersed in a directional dc electric field region. Nevertheless, the EK phenomena (e.g., dielectrophoresis, electrophoresis, and electro-osmosis) have been widely employed to induce mechanical motion on microscale and nanoscale entities in the areas of biotechnology (e.g., cell trapping and manipulation) and, more recently, in nanoelectronics applications (e.g., carbon nanotube and inorganic nanowire manipulation). In this section, the basic principles of these three phenomena will be discussed, followed by the examples showing how researchers worldwide employ the EK phenomena to manipulate nanoscale entities for nanoelectronics and biotechnology applications.

4.1. Basic Principles

4.1.1. Dielectrophoresis

Assume a system consisting of a neutral colloidal particle (with complex permittivity ε_1) inside a fluid medium (with complex permittivity ε_2); applying a directional electric field will cause the polarization of both the colloidal particle and the fluid medium. Consider the simplest case where the fluid medium is a vacuum space. Applying a directional electric field will only cause the colloidal particle to polarize. The polarized particle will tend to align itself with the directional electric field which results in a dipole moment ($\vec{P}$), and Coulombic forces will then be induced on the polarized particle by the electric field ($\vec{E}$). If an oscillating and nonuniform electric field is applied instead (i.e., ac heterogeneous electric field), then a gradient of electric field exists which results in an imbalance of Coulombic forces acting on both sides of the particle. Hence the resultant force acting on the particle will drive it into motion (see Fig. 5). The resultant force acting on the particle is termed dielectrophoretic force ($\vec{F}_{\text{DEP}}$). This force can be expressed mathematically as [113]

$$\vec{F}_{\text{DEP}} = (\vec{P} \cdot \nabla)\vec{E} \tag{1}$$

where ∇ is the gradient operator. For a particular case where the colloidal particle is isotropically, linearly, and homogeneously polarizable, the dipole moment induced can be expressed as

$$\vec{P} = \alpha V \vec{E} \tag{2}$$

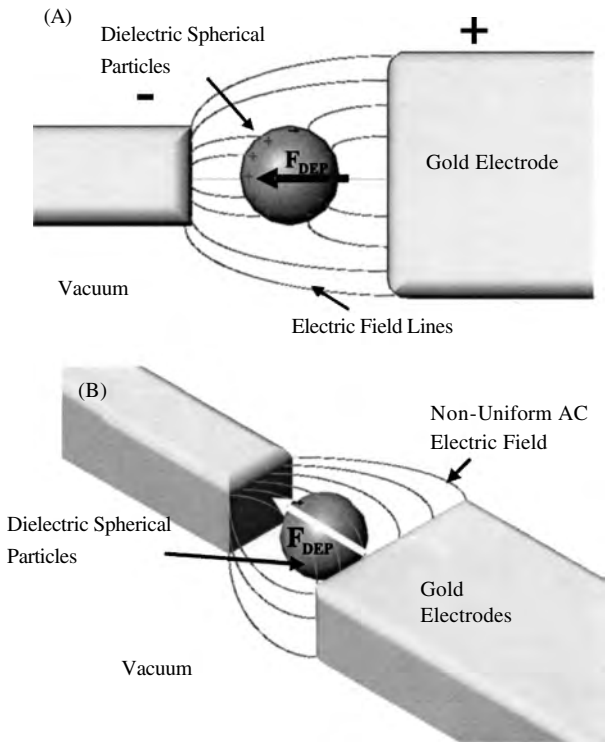


Figure 5. Principle of dielectrophoresis. (A) A spherical dielectric particle is polarized under the influence of external electric field at time Δt. (B) Under nonuniform ac electric field, dielectrophoretic force induced on the neutral spherical particle causes the mechanical motion.

where α is the polarizability and V is the volume of the particle. Combining Eq. (1) and (2), the final expression of dielectrophoretic force can be expressed as

$$\vec{F}_{\text{DEP}} = \frac{1}{2}\alpha V \nabla |\vec{E}|^2 \tag{3}$$

For general cases, if the fluid medium is not a vacuum space, then the dielectric constants for both the particle and the fluid medium have to be taken into consideration. For a spherical colloidal particle (with radius r), Eq. (3) can be further expressed in terms of the complex permittivities of the particle and the medium,

$$\vec{F}_{\text{DEP}} = 2\pi r^3 \varepsilon_2 \,\text{Re}[K(\omega)] \nabla |\vec{E}|^2 \tag{4}$$

where $\text{Re}[K(\omega)]$ is the real part of $K(\omega)$, which is defined by $(\varepsilon_1 - \varepsilon_2)/(\varepsilon_1 + 2\varepsilon_2)$. $K(\omega)$ is referred to as the Clausius–Mossotti factor, which is an important quantity to determine the direction of the dielectrophoretic force acting on a particle. The Clausius–Mossotti factor is a frequency dependent term due to the frequency dependency of the complex permittivities of materials.

From Eq. (4), there are two important points about dielectrophoresis. First, the process of dielectrophoresis depends on the gradient of the electric field strength instead of the direction of the field. Second, if the fluid medium is more polarizable than that of the colloidal particle (i.e., $\varepsilon_2 > \varepsilon_1$), then the dielectrophoretic force induced will be negative. In other words, the colloidal particle will be repelled from the region of high electric field to the region of low electric field (i.e., negative dielectrophoresis). Similarly, positive dielectrophoresis results when the colloidal particle is more polarizable than that of the fluid medium (i.e., $\varepsilon_1 > \varepsilon_2$).

4.1.2. Electrophoresis

Unlike dielectrophoresis, which can induce motion on neutral dielectric particles under a nonuniform ac electric field, electrophoresis is only effective on charged particles. Under a directional electric field generated by metal electrodes, charged particles will move toward the electrodes which are oppositely charged (see Fig. 6). In order to give insight on how electrophoresis applies on charged particles, a simple example is considered here. Assume a spherical particle of radius r, charge q, and mass m, which is in a fluid of viscosity η in a one-dimensional space system under a homogenous dc electric field E. The external electric field E will exert a Coulombic force ($F_{\text{comb}} = qE$) on the particle and drive the particle into motion. However, at the same time, a viscous drag force (or Stokes' force, $F_{\text{drag}} = -6\pi r \eta \dot{x}$) will develop to counteract the motion of the particle. Therefore the net resultant force acting on the particle is given by Newton's second law,

$$m\ddot{x} = qE - 6\pi r \eta \dot{x} \tag{5}$$

Solving Eq. (5), the equation of motion of the particle is expressed in the following forms, assuming the particle is initially at rest:

$$x = \frac{mqE}{36\pi^2 r^2 \eta^2}\left(e^{\frac{-6\pi r\eta}{m}t} + \frac{6\pi r \eta}{m}t - 1\right) \tag{6}$$

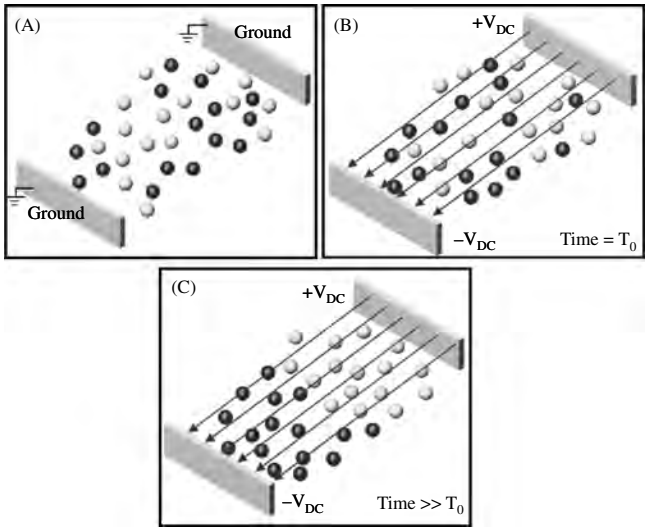


Figure 6. Principle of electrophoresis. Charged particle (orange: positively charged; blue: negatively charged) will migrate toward its opposite polarity under the influence of external electric field. (A) Mixture of charged particles with no external electric field. (B) Application of external electric field imparts forces on the charged particles; the particles move toward either the cathode or anode, depending on the net charge the particle possesses. (C) Separation of particles with different net charges will eventually be achieved.

$$\dot{x} = \frac{qE}{6\pi r\eta}(1 - e^{\frac{-6\pi r\eta}{m}t}) \tag{7}$$

From Eq. (7), the velocity of the particle will eventually come to a constant value after a finite amount of time elapse (i.e., $\dot{x} = \frac{qE}{6\pi r\eta}$), which means that the forces acting on the body are balanced, and results in constant velocity motion (i.e., Newton's first law). This is only the case when the particle is immersed in an environment where the electric field and viscosity of the medium are homogenous. As implied in Eq. (7), the velocity of the particles depends on the net charge of the particles, and therefore, electrophoresis does provide a powerful tool for separating particles with different net charges.

4.1.3. Electro-osmosis

In the simplest context, electro-osmosis is the movement of charged ions (which are formed near the surface of the solid–liquid interface) by an external electric field, which in turn moves the solvent surrounding the ions by viscous friction (or other physical forces like hydrogen bonding of water molecules or van der Waals forces, etc.), and resulting in macroscopic fluid movement (see Fig. 7). The most important factor that contributes to electro-osmosis is the formation of an electrical double layer (EDL) (i.e., a layer of charged ions) in the interface between the solid–liquid boundaries. An EDL is a layer consisting of either majority in cation or anion depending on the polarity of the charged surface. There are two possible mechanisms for the formation of an EDL [114]. The first proposed mechanism is due to surface dissociation (i.e., protons leave the surface and hence become negatively charged). For example, when silica is in contact with an aqueous solution, its surface hydrolyzes

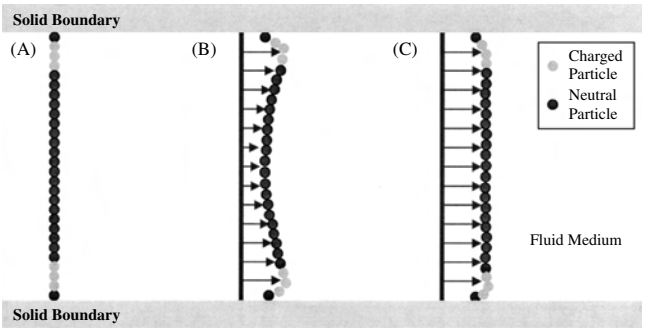


Figure 7. Original electro-osmosis mechanism proposed by Quincke and Holmholtz. The system consists of solid boundary and liquid molecules (green: charged molecules; orange: uncharged molecules). (A) Without the application of external electric field. (B) Application of external electric field causes the charged molecules (green) to move, and the viscous drag forces are then acting upon on the uncharged particles. (B) Electric field is applied for a period of time and the flow of the liquid becomes steady (i.e., constant velocity profile along the envelope).

to form silanol surface group which may be either positively charged ($SiOH_2^+$), neutral (SiOH), or negatively charged (SiO^-) [115]. The second mechanism is the bounding of the ions that already exist in the liquid. The length of the EDL is given by the reciprocal of the Debye function. Interested readers can refer to Tikhomolova [116] for more information on the Debye function.

To have a fundamental understanding between the fluid motion and the applied electric field, assume a scenario of liquid motion inside a capillary tube with sufficient large cross sectional area so that the EDLs of the solution are not overlapping. The velocity of the electro-osmotic flow (v_{EOF}) is then directly proportional to the strength of the electric field strength (E) and it was termed quantitatively by Smoluchowski in 1903 as [116]

$$v_{EOF} = \frac{\varepsilon\xi}{\eta}E \tag{8}$$

where ξ is the zeta potential (i.e., the potential between the charge surface and the electrolyte solution), ε is the dielectric constant of the fluid medium, and η is the viscosity of the solution.

4.2. Electrokinetic Manipulation of Carbon Nanotube, Metallic, and Inorganic Nanowires

With ever-increasing demand in the microelectronics industry, scientists and engineers have pushed the manufacturing limit of the minimum feature on a chip to submicrometer range in the late 20th century. Currently, the minimum feature of a typical Intel® Pentium® 4 CPU has already reached 0.13 μm [117]. However, the continual scaling down of the microelectronic devices will become physically and economically impractical in the foreseeable future. Therefore, scientists have to look into new materials and processing technologies to create *nanoelectronics* for the future.

In 1991, Iijima discovered the carbon nanotube [4], which has opened up a new research area for nanoelectronic applications. Carbon nanotube based nanoelectronic devices have already been fabricated successfully in the past few years. For example, Tans et al. [118] and Frank et al. [119] have demonstrated the single carbon nanotube transistor and quantum resistors, respectively. On another ocassion, Duan et al. exploited the usage of indium phosphide nanowires for nanoscale electronic and optoelectronics components [120]. Although many promising results have been demonstrated in using carbon nanotubes and nanowires for nanoelectronic applications, techniques to batch manipulate (or parallel assemble) those nanoentities have to be developed in order to make nanoelectronics feasible and practical. Currently, manipulation of carbon nanotube and nanowires by electrokinetic forces is one of the most active research areas for parallel assembly techniques.

Using existing micro/nanolithography technology, metal electrodes with dimensions in the range of a nanometer are currently achievable [121]. With smaller electrode pairs, ultrahigh electric field strength ($>10^6$ V/m) can be produced with low applied voltages. For example, consider a pair of microelectrodes with a gap distance of 10 μm; an electric field strength of 10^6 V/m can be produced easily by an excitation voltage of 10 V across the microelectrodes. This intense electric field strength is useful for micro/nanomanipulation.

Several groups have demonstrated successful manipulation of carbon nanotube, metallic, and inorganic nanowires using electrokinetic forces in the past few years. Yamamoto et al. have demonstrated the use of electrophoretic force to align carbon nanotubes with the dc electric field [122]. Later, the same group exploited the use of dielectrophoretic forces to separate the carbon nanotubes and the impurities which serve as a novel purification process [123]. Recently, Nagahara et al. have successfully demonstrated the use of dielectrophoretic force to manipulate individual carbon nanotubes to form electrical contacts between two metal nanoelectrodes, which is a very promising demonstration for parallel assembly of carbon nanotubes [121]. Wong and Li have recently demonstrated the formation of bulk carbon nanotubes between microelectrodes (see Fig. 8) using dielectrophoretic forces for practical sensing applications [124]. Other than carbon nanotubes, metallic and inorganic nanowires have also been manipulated successfully by Smith et al. [125] and Duan et al. [120], respectively. In particular, Dual et al. have demonstrated the parallel assembly of inorganic nanowires via an array of parallel electrode pairs, indicating an important step toward batch fabrication of nanoelectronic elements [120].

4.3. Electrokinetic Manipulation of Biological Entities and Nanoparticles

An early attempt at manipulating biological entities such as DNA molecules was shown successfully by Washiuzu and Kurosawa [126] in 1990. Many exciting results in manipulating biological entities and nanoparticles using electrokinetic approaches have been published since then.

In order to effectively investigate the effect of dielectrophoretic force on biological entities and nanoparticles,

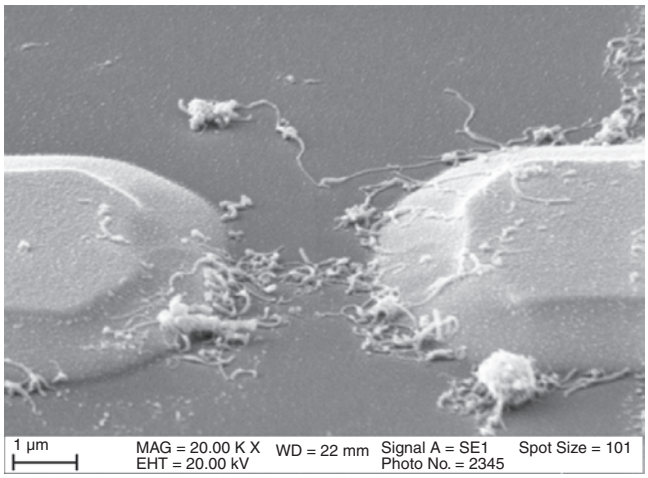


Figure 8. Scanning electron microscopic image showing the bridging of Au/Cr microelectrodes by bundled carbon nanotube by dielectrophoretic force.

Huang and Pethig developed a simple and practical methodology in designing electrode geometries in producing well-defined nonuniform electrode fields in the early 1990s [127], which was shown to be useful in collecting yeast cells to or away from the electrodes. Practical application in using dielectrophoretic force to separate leukaemic cells from human blood was demonstrated by Becker et al. in the mid 1990s [128]. In late 1990s, Hughes and Morgan successfully demonstrated the trapping of single viruses, capsids, and latex spheres, with diameters in the range of 93–250 nm inside a potential trap [129]. By utilizing both positive and negative dielectrophoretic forces, precise movement of protoplasts with 30–50 μm diameters can be controlled by a three-dimensional grid system developed by Suehiro and Pethig [130]. As implied by Eq. (4), the direction and magnitude of the dielectrophoretic force are dependent on the dielectric properties of the particles and medium, and also the electric field gradient in the medium. Using this property, a system of particles with different dielectric properties can be separated into subpopulations by dielectrophoresis and this has been successfully demonstrated by Green and Morgan in separating a carboxylate modified latex bean of 93 nm diameter [131]. Besides trapping the particles in a certain predefined space, dielectrophoresis has been found useful to mix a system of particles in chaotic mixing applications for microfluidic systems as well [132]. Apart from the dielectrophoretic approach in manipulating the biological entities, the electrophoretic and electro-osmotic approach has been proven to work in manipulating those particles as well. For instance, a conductive nanoparticle (Pd) of about 17 nm diameter can be trapped electrostatically by nanoelectrodes as shown by Bezryadin et al. [133].

5. CONCLUSION

As described in this chapter, the three most widely accepted techniques for robotic manipulation of nanoentities are: (1) manipulation by atomic force microscope tips, (2) manipulation by laser trapping force, and (3) manipulation by electrokinetic forces. However, besides chemical and

biotechnology methods for nanomanipulation, which are not described in this chapter, there are a few other techniques which have been demonstrated for nanomanipulation. For example, in terms of single molecule manipulation, magnetic particles can be attached to the molecules and then magnetic force can be used to move the molecules [134]. Hydrodynamic forces (i.e., shear force) can also be used to move or stretch molecules [131, 135, 136], and ultrasmall glass needles have also been demonstrated to manipulate molecules [137]. Nevertheless, none of the existing nanorobotic and manipulation techniques is able to solve the most pressing problem if true nanorobots are to be created some day: the ability to bring two or more molecules and join them together in a desired manner to create the first controllable "nanotool"—a tool that can then be used to make more complex nanosystems such as intelligent nanorobots that are able to autonomously manufacture to more nanosystems.

GLOSSARY

Atomic force microscopy (AFM) A form of scanning probe microscopy. The tip–sample interaction monitored during raster scan is interatomic forces between the tip and sample surface. During scanning, the cantilever deflection or interatomic force between tip and sample is kept constant in contact mode and the rms value of cantilever deflection is kept constant in TappingMode™.

Dielectrophoresis The phenomenon of inducing motion on neutral dielectric particles under nonuniform ac electric field.

Dielectrophoretic force The force generated when a neutral particle is immersed in nonuniform ac electric field region.

Electrokinetic force The Columbic force imparted on particles when external electric field is applied.

Electro-osmosis The phenomenon of moving of charged ions, formed near the solid–liquid interface, by an external electric field. The solvent that surrounds the ions is also moved by viscous friction or other physical forces, like hydrogen bonding of water molecules or van der Waals forces, resulting in macroscopic fluid movement.

Electrophoresis The phenomenon of inducing motion on charged particles under directional dc electric field.

Electrophoretic force The force generated when a charged particle is immersed in a directional dc electric field region.

Haptic device A device that allows position and orientation input from and forces/torques output to operators at the same time.

Mie regime Optical trapping occurs when the wavelength of the laser is greatly smaller than the object size.

Nanoassembly Study of how to fit together a group of nanometer sized components to form a functional machine or device.

Nanolithography The process of imprinting patterns of submicrometer size on a sample surface.

Nanomanipulation Study on how to manipulate nanometer sized entities.

Nanorobotics Study of dynamics, control, planning, sensing, and actuation of robots in nanometer scale.

Optical trapping The phenomenon of fixing or holding micrometer or submicrometer sized neutrally charged entities using focused laser. The trapping ability comes from the radiation pressure of light, which is generated by the change of momentum in the transmitted and incident laser light and is balanced by all external forces, like gravity and electric forces.

Optical tweezers A laser based device for manipulating micrometer or submicrometer sized neutral entities based on the phenomenon of laser trapping.

Optical tweezing See optical trapping.

Piezoelectric scanner A core component of scanning probe microscopes. An $X/Y/Z$ scanner that is actuated by three pairs of piezoelectric materials. Piezoelectric material extends or contracts when external ac electric field is applied.

Position sensitive detector (PSD) A core component of scanning probe microscopes. It consists of two or four photodiodes to detect the minute changes in the light path of a laser which reveal the deflection and twist of the cantilever in AFM and STM.

Rayleigh regime Optical trapping occurs when the wavelength of the laser is comparable to the object size.

Scanning probe microscopy (SPM) A family of microscopy techniques that generate 3D topography and images of various physical properties, for example height, surface roughness, local resistivity, etc., of sample surfaces with nanometer or subnanometer scale. A sharp tip is rastered across the sample surface while tip–sample interactions are monitored.

Scanning tunneling microscopy (STM) A form of scanning probe microscopy. The tip–sample interaction monitored during raster scan is tunneling current between the tip and sample surface. Since the tunneling current has an exponential relationship with tip–sample separation, the height information is derived directly from the measured tunneling current in the generation of 3D topography.

Virtual reality (VR) The technique of computer simulations that use 3D graphics and special input devices to allow users to interact with the simulation.

REFERENCES

1. K. E. Drexler, "Engines of Creation." Anchor, 1986. Available online at: www.foresight.org/EOC/index.html.
2. K. E. Drexler, "Nanosystems: Molecular Machinery, Manufacturing and Computation." Wiley, New York, 1992.
3. A. M. Morales and C. M. Lieber, *Science* 279, 208 (1998).
4. S. Iijima, *Nature* 354, 56 (1991).
5. C. J. Murphy, *Science* 298, 2139 (2002).
6. A. J. Heinrich, C. P. Lutz, J. A. Gupta, and D. M. Eigler, *Science* 298, 1381 (2002).
7. D. M. Eigler and E. K. Schweizer, *Nature* 344, 524 (1990).
8. P. Kim and C. M. Lieber, *Science* 286, 2148 (1999).
9. R. S. Fearing, in "Proc. IEEE/RSJ Int'l Conf. Intelligent Robots and Systems 1995," Vol. 2, pp. 212–217.
10. M. Sitti, in "Proc. IEEE Int'l Conf. Nanotechnology," 2001, pp. 75–80.
11. A. M. Fennimore, T. D. Yuzvinsky, W. Q. Han, M. S. Fuhrer, J. Cumings, and A. Zettl, *Nature* 424, 408 (2003).

12. R. K. Soong, G. D. Bachand, H. P. Neves, A. G. Olkhovets, H. G. Craighead, and C. D. Montemagno, *Science* 290, 1555 (2000).
13. A. A. G. Requicha, *Proc. IEEE* 11, 1928 (2003).
14. E. Bonabeau, M. Dorigo, and G. Theralaz, "Swarm Intelligence: From Natural to Artificial Systems." Oxford University Press, Oxford, 1999.
15. A. Cavalcanti, *IEEE Trans. Nanotechnol.* 2, 82 (2003).
16. Biotech as the Fastest Pathway to an Assembler, Molecubotics, Inc., 2001. Available online at: www.molecubotics.com/tech-docs/biotech-pathway.html.
17. J. A. Liddle, L. R. Harriott, A. E. Novembre, and W. K. Waskiewicz, SCALPEL: A Projection Electron-Beam Approach to Sub-Optical Lithography, Tech. Rev., Bell Labs, Lucent Technologies, 1999. Available online at: www.bell-labs.com/project/SCALPEL.
18. A. Katz and M. E. Davis, *Nature* 403, 286 (2000).
19. R. D. Piner, J. Zhu, F. Xu, S. Hong, and C. A. Mirkin, *Science* 283, 661 (1999).
20. G. M. Whitesides and B. Grzybowski, *Science* 295, 2418 (2002).
21. K. Tanaka, A. Tengeiji, T. Kato, N. Toyama, and M. Shionoya, *Science* 299, 1212 (2003).
22. D. A. Bonnell, Eds., Scanning Probe Microscopy and Spectroscopy: Theory, Techniques and Applications, John Wiley & Sons (2000).
23. G. Binning, C. F. Quate, and Ch. Gerbre, *Phys. Rev. Lett.* 56, 930 (1986).
24. R. S. Becker, J. A. Golovchenko, and B. S. Swartzentruber, *Nature* 325, 419 (1987).
25. H. J. Mamin, P. H. Guethner, and D. Rugar, *Phys. Rev. Lett.* 65, 2418 (1990).
26. H. J. Mamin, S. Chiang, H. Birk, P. H. Guethner, and D. Rugar, *J. Vac. Sci. Technol. B* 9, 1398 (1991).
27. D. Clery, *New Scientist* 129, 31 (1991).
28. L. J. Whitman, J. A. Stroscio, R. A. Dragoset, and R. J. Celotta, *Science* 251, 1206 (1991).
29. M. Aono, *Science* 258, 586 (1992).
30. O. M. Leung and M. C. Goh, *Science* 255, 64 (1992).
31. J. A. Stroscio and D. M. Eigler, *Science* 254, 1319 (1991).
32. M. F. Crommie, C. P. Lutz, and D. M. Eigler, *Science* 262, 218 (1993).
33. G. Dujardin, R. E. Walkup, and P. Avouris, *Science* 255, 1232 (1992).
34. Y. W. Mo, *Science* 261, 886 (1993).
35. M. Sitti, S. Horighuchi, and H. Hashimoto, in "Proceedings of 1999 IEEE/RSJ International Conference on Intelligent Robots and Systems," Vol. 2, pp. 882–888.
36. M. Guthold, M. R. Falvo, W. G. Matthews, S. Paulson, S. Washburn, D. A. Erie, R. Superfine, F. P. Brooks, Jr., and R. M. Taylor II, *IEEE/ASME Trans. Mechatron.* 5, 89 (2000).
37. L. Dong, F. Arai, and T. Fukuda, in "Proceedings of 2001 IEEE International Conference on Robotics and Automation," Seoul, South Korea, Vol. 1, pp. 632–637.
38. M. F. Yu, M. J. Dyer, G. D. Skidmore, H. W. Rohrs, X. K. Lu, K. D. Ausman, J. R. von Eher, and R. S. Ruoff, *Nanotechnology* 10, 244 (1999).
39. S. Fahlbusch, A. Shirinov, and S. Fatikow, in "Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems," Vol. 3, pp. 1772–1777.
40. E. Hæggström, G. G. Yaralioglu, A. S. Ergun, and P. T. Khuri-Yakub, in "Proceedings of the 2002 Second International Conference on Nanotechnology," pp. 45–49.
41. L. Dong, F. Arai, and T. Fukuda, in "Proceedings of the 2002 Second International Conference on Nanotechnology," Washington, DC, pp. 443–446.
42. B. C. Yao, F. S.-S. Chien, S. Chen, P. K. W. Lui, and G. S. Peng, in "Proceedings of the 2002 Second International Conference on Nanotechnology," pp. 13–15.
43. G. Li, N. Xi, M. Yu, and W. K. Fung, in "Proc. of the 2003 International Conference on Robotics and Automation," pp. 3642–3647.
44. J. M. Friedt, M. Hoummady, and J. Cervelle, in "Proceedings of the 1999 IEEE/ASME International Conference on Advanced Intelligent Mechatronics," pp. 9–12.
45. Y. Kim and C. M. Lieber, *Science* 257, 5068 (1992).
46. R. Luthi, E. Meyer, H. Haefke, L. Howald, W. Gutmannsbauer, and H. J. Guntherodt, *Science* 266, 5193 (1994).
47. D. M. Schaefer, R. Reifenberger, A. Patil, and R. P. Andres, *Appl. Phys. Lett.* 66, 1012 (1995).
48. T. Junno, K. Deppert, L. Montelius, and L. Samuelson, *Appl. Phys. Lett.* 66, 3627 (1995).
49. A. A. G. Requicha, C. Baur, A. Bugacov, B. C. Gazen, B. Koel, A. Madhukar, T. R. Ramachandran, R. Resch, and P. Will, in "Proceedings of the IEEE International Conference on Robotics and Automation," Leuven, Belgium, 1998, pp. 3368–3374.
50. R. Resch, D. Lewis, S. Meltzer, N. Montoya, B. E. Koel, A. Madhukar, A. A. G. Requicha, and P. Will, *Ultramicroscopy* 82, 135 (2000).
51. A. A. G. Requicha, S. Meltzer, F. P. T. Arce, J. H. Makaliwe, H. Siken, S. Hsieh, D. Lewis, B. E. Koel, and M. Thompson, in "Proceedings of the First IEEE International Conference on Nanotechnology," Maui, HI, 28–30 October 2001, pp. 81–86.
52. S. Hsieh, S. Meltzer, C. R. C. Wang, A. A. G. Requicha, M. E. Thompson, and B. E. Koel, *J. Phys. Chem. B* 106, 231 (2002).
53. T. Hertel, R. Martel, and P. Avouris, *J. Phys. Chem. B* 102, 910 (1998).
54. H. W. Ch. Postma, M. de Jonge, Z. Yao, and C. Dekker, *Phys. Rev. B* 62, 653 (2000).
55. M. R. Falvo, M. R. Taylor II, A. Hesler, V. Chi, F. P. Brooks, Jr., S. Washburn, and R. Superfine, *Nature* 397, 236 (1999).
56. M. R. Falvo, G. J. Clary, R. M. Taylor II, V. Chi, F. P. Brooks, Jr., S. Washburn, and R. Superfine, *Nature* 389, 6651 (1997).
57. T. W. Tombler, C. Zhou, L. Alexseyev, J. Kong, H. Dai, L. Liu, C. S. Jayanthi, M. Tang, and S. Y. Wu, *Nature* 405, 769 (2000).
58. L. Dong, F. Arai, and T. Fukuda, in "Proceedings of 2002 IEEE International Conference on Robotics and Automation," Washington, DC, Vol. 2, pp. 1477–1482.
59. L. T. Hansen, A. Kuhle, A. H. Sorensen, J. Bohr, and P. E. Lindelof, *Nanotechnology* 9, 337 (1998).
60. M. Sitti and H. Hashimoto, in "Proceedings of 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems," Vol. 3, pp. 1739–1746.
61. V. J. Morris, A. P. Gunnung, and A. R. Kirty, "Atomic Force Microscopy for Biologists." Imperial College Press, 1999.
62. "Atomic Force Microscopy in Cell Biology (Methods in Cell Biology)" (B. P. Jena and J. K. Heinrich Horber, Eds.). Academic Press, San Diego, 2002.
63. M. Gutholdm, W. G. Matthews, A. Negishi, R. M. Taylor II, D. Erie, F. P. Brooks, Jr., and R. Superfine, *Surf. Interf. Anal.* 27, 437 (1999).
64. J. H. Hoh, R. Lal, S. A. John, J.-P. Revel, and M. F. Arnsdof, *Science* 253, 1405 (1991).
65. H. G. Hansma, J. Vesenka, C. Siegerist, G. Kelderman, H. Morrett, R. L. Sinsheimer, V. Elings, C. Bustamante, and P. K. Hansma, *Science* 256, 1190 (1992).
66. S. Thalhammer, R. W. Stark, S. Müller, J. Wienberg, and W. M. Heckl, *J. Structural Biol.* 119, 232 (1997).
67. R. W. Stark, S. Thalhammer, J. Wienberg, and W. M. Heckl, *Appl. Phys. A* 66, 579 (1998).
68. K. E. Drexler and J. S. Foster, *Nature* 343, 600 (1990).
69. E. L. Florin, V. T. Moy, and H. E. Gaub, *Science* 264, 415 (1994).
70. U. Dammer, M. Hegner, D. Anselmetti, P. Wagner, M. Dreier, W. Huber, and H.-J. Güntherodt, *Biophys. J.* 70, 2437 (1996).
71. R. Ros, F. Schwesinger, D. Anselmetti, M. Kubon, R. Schäfer, A. Plückthun, and L. Tiefenaufer, *Proc. Natl. Acad. Sci. Biophys.* 95, 7402 (1998).

72. A. Raab, W. Han, D. Badt, S. J. Smith-Gill, S. M. Lindsay, H. Schindler, and P. Hinterdorfer, *Nature Biotechnol.* 17, 902 (1999).
73. M. Sitti, in "Proceedings of the First International Conference on Nanotechnology," Maui, HI, 28–30 October 2001, pp. 99–104.
74. A. Ashkin, *Phys. Rev. Lett.* 24, 156 (1970).
75. A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm, and S. Chu, *Opt. Lett.* 11, 288 (1986).
76. A. Ashkin, *Biophys. J.* 61, 2, 2569 (1992).
77. "Laser Tweezers in Cell Biology" (M. P. Sheetz, L. Wilson, and P. Matsudaira, Eds.). Academic Press, San Diego, 1998.
78. A. Ashkin and J. M. Dziedzic, *Science* 235, 1517 (1987).
79. A. Ashkin, J. M. Dziedzic, and T. M. Yamane, *Nature* 330, 769 (1987).
80. K. J. Aufderheide, Q. Du, and E. S. Fry, *J. Eukary. Microbiol.* 40, 793 (1993).
81. J. M. Colon, P. G. Sarosi, and P. G. McGovern, *Fertility and Sterility* 57, 695 (1992).
82. Y. Tadir, W. H. Wright, O. Vafa, T. Ord, R. Asch, and M. W. Berns, *Fertility and Sterility* 52, 870 (1989).
83. Y. Tadir, W. H. Wright, O. Vafa, T. Ord, L. H. Liaw, R. Asch, and M. W. Berns, *Human Reproduction* 6, 1011 (1991).
84. J. Conia and S. Voelkel, *BioTechniques* 17, 1162 (1994).
85. K. Schütze, A. Clement-Sengewald, and A. Ashkin, *Fertility and Sterility* 61, 783 (1994).
86. A. Clement-Sengewald, K. Schütze, A. Ashkin, G. A. Palma, G. Kerlen, and G. Brem, *J. Assist. Reprod. Genet.* 13, 259 (1996).
87. Y. Tadir, J. Neev, and M. W. Berns, *Seminar Reproductive Endocinol.* 12, 169 (1994).
88. G. B. Blanchet, J. B. Russell, C. R. Fincher, Jr., and M. Portmann, *Fertility and Sterility* 57, 1337 (1992).
89. G. Weber and K. O. Greulich, *Int. Rev. Cytol.* 133, 1 (1992).
90. G. Weber, S. Monajembashi, K. O. Greulich, and J. Wolfrum, *Eur. J. Cell. Biol.* 49, 37 (1989).
91. G. Weber, A. Weith, R. Kaiser, K. H. Grzeschik, and K. Olek, *Somat. Cell. Mol. Genet.* 16, 123 (1990).
92. H. Liang, W. H. Wright, W. He, and M. W. Berns, *Exp. Cell Res.* 204, 110 (1993).
93. H. Liang, W. H. Wright, C. L. Rieder, E. D. Salmon, G. Profeta, J. Andrews, Y. Liu, G. J. Sonek, and M. W. Berns, *Exp. Cell Res.* 213, 308 (1993).
94. S. Bayoudh, M. Mehta, H. Rubinsztein-Dunlop, N. R. Heckenberg, and C. Critchley, *J. Microsc.* 203, 214 (2001).
95. K. Sakata-Sogawa, M. Kurachi, K. Sogawa, Y. Fujii-Kuriyama, and H. Tashiro, *Eur. Biophys. J.* 27, 55 (1998).
96. P. J. Bronkhorst, J. Grimbergen, G. J. Brackenhoff, R. M. Heethaar, and J. J. Sixma, *Br. J. Haematol.* 96, 256 (1997).
97. C. Viegal, M. Bartoo, D. C. S. White, J. C. Sparrow, and J. E. Molley, *Biophys. J.* 75, 1424 (1998).
98. M. L. Bennink, O. D. Scharer, and R. Kanaar, *Cytometry* 36, 200 (1999).
99. T. N. Buican, M. J. Smyth, H. A. Crissman, G. C. Salzman, C. C. Stewart, and J. C. Martin, *Appl. Opt.* 26, 5311 (1987).
100. J. G. Mitchell, R. Weller, and M. Beconi, *Microbial Ecol.* 25, 113 (1993).
101. R. Huber, S. Burhhraf, T. Mayer, S. M. Barns, P. Rossnagle, and K. O. Stetter, *Nature* 376, 57 (1996).
102. A. Ashkin, *Nature* 330, 608 (1987).
103. S. Chu, *Science* 252, 861 (1991).
104. A. Ashkin, *Proc. Natl. Acad. Sci. Phys.* 94, 4853 (1997).
105. H. Furkawa and I. Yamaguchi, *Opt. Lett.* 23, 216 (1998).
106. K. Svoboda and S. M. Block, *Opt. Lett.* 19, 930 (1994).
107. M. E. J. Friese, T. A. Nieminen, N. R. Heckenberg, and H. Rubinszterin-Dunlop, *Nature* 394, 348 (1998).
108. L. Paterson, M. P. MacDonald, J. Arlt, W. Sibbett, P. E. Bryant, and K. Dholakia, *Science* 292, 912 (2001).
109. M. P. MacDonald, L. Paterson, K. Volke-Sepulveda, J. Arlt, W. Sibbett, and K. Dholakia, *Science* 296, 1101 (2002).
110. S. A. Diddams, Th. Udem, J. C. Berquist, E. A. Curtis, R. E. Drullinger, L. Hollberg, W. M. Itanom W. D. Lee, C. W. Oates, K. R. Vogel, and D. J. Wineline, *Science* 293, 825 (2001).
111. L. M. Duan, J. I. Cirac, and P. Zoller, *Science* 292, 1695 (2001).
112. D. Kielpinski, V. Meyer, M. A. Rowe, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Winelind, *Science* 291, 1013 (2001).
113. H. A. Pohl, "Dielectrophoresis: The Behavior of Neutral Matter in Nonuniform Electric Fields." Cambridge Univ. Press, Cambridge, UK, 1978.
114. C. M. Ho, in "Proc. IEEE Int'l Conf. MEMS," 2001, pp. 375–384.
115. R. Weinberger, "Practical Capillary Electrophoresis." Academic Press, San Diego, 1993.
116. K. P. Tikhomolova, "Electro-osmosis." Ellis Horwood, New York, 1993.
117. Intel® Pentium® 4 Processor Datasheet, 2002. Available online at: www.intel.com.
118. S. J. Tans, A. R. M. Verschueren, and C. Dekker, *Nature* 393, 49 (1998).
119. S. Frank, P. Poncharal, Z. L. Wang, and W. A. de Heer, *Science* 280, 1744 (1998).
120. X. Duan, Y. Huang, Y. Cui, J. Wang, and C. M. Lieber, *Nature* 409, 66 (2001).
121. L. A. Nagahara, I. Amlani, J. Lewenstein, and R. K. Tsui, *Appl. Phys. Lett.* 80, 3826 (2002).
122. K. Yamamoto, S. Akita, and Y. Nakayama, *Japan. J. Appl. Phys.* 35, L917 (1996).
123. K. Yamamoto, S. Akita, and Y. Nakayama, *J. Phys. D: Appl. Phys.* 31, L34 (1998).
124. V. T. S. Wong and W. J. Li, in "Proc. IEEE Int'l Conf. MEMS," 2003, pp. 41–44.
125. P. A. Smith, C. D. Nordquist, T. N. Jackson, T. S. Mayer, B. R. Martin, J. Mbindyo, and T. E. Mallouk, *Appl. Phys. Lett.* 77, 1399 (2000).
126. M. Washizu and O. Kurosawa, *IEEE Trans. Ind. App.* 26, 1165 (1990).
127. Y. Huang and R. Pethig, *Meas. Sci. Technol.* 2, 1142 (1991).
128. F. F. Becker, X. B. Wang, Y. Huang, R. Pethig, J. Vykoukal, and P. R. C. Gascoyne, *J. Phys. D: Appl. Phys.* 27, 2659 (1994).
129. M. P. Hughes and H. Morgan, *J. Phys. D: Appl. Phys.* 31, 2205 (1998).
130. J. Suehiro and R. Pethig, *J. Phys. D: Appl. Phys.* 31, 3298 (1998).
131. N. G. Green and H. Morgan, *J. Phys. D: Appl. Phys.* 30, L41 (1997).
132. J. Deval, P. Tabeling, and C. M. Ho, in "Proc. IEEE Int'l Conf. MEMS," 2002, pp. 36-39.
133. A. Bezryadin, C. Dekker, and G. Schmid, *Appl. Phys. Lett.* 71, 1273 (1997).
134. S. B. Smith, L. Finzi, and C. Bustamante, *Science* 258, 1122 (1992).
135. T. T. Perkins, D. E. Smith, R. G. Larson, and S. Chu, *Science* 268, 83 (1995).
136. P. K. Wong, Y. K. Lee, and C. M. Ho, *J. Fluid Mech.* 497, 55 (2003).
137. S. B. Smith, Y. Cui, and C. Bustamante, *Science* 271, 795 (1996).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanorecognition

P. Tarakeshwar, Kwang S. Kim

Pohang University of Science and Technology, Pohang, Koreá

CONTENTS

1. INTRODUCTION

The motivation to comprehend the underlying principles of nanorecognition stems from our need to understand, predict, and design the properties of nanoscale materials and devices with respect to their chemical composition, physical characteristics, and electronic structures [1–35]. Nanorecognition in many ways is synonymous to the more well-known "molecular recognition" [36–95], but quantum effects are predominant [96–135]. For example, the recognition of (single) electrons and photons by molecules or clusters of atoms, which are of interest in the design and working of nanodevices, can only be described by using a model which explicitly takes into account the quantum characteristics of electrons and photons.

The phenomena of recognition is widespread in most chemical and biological systems and in fact is the key to the understanding of biological machinery. It is the basis of the specificity observed in enzyme–substrate [136–148], hormone–receptor [149–153], antigen–antibody [154–159], and drug–receptor interactions [160–168], in biological systems. It also leads to protein folding, oligomerization, aggregation, and the formation of nucleic acid helices [169–188]. In fact, much of the inspiration for the developments in contemporary supramolecular and host–guest chemistry emerged from efforts to understand and mimic the exquisite and intricate design observed in biological systems.

Given its widespread relevance in diverse fields of chemistry, biology, and more recently materials science, there have been intense efforts to obtain a detailed understanding of the underlying basis of molecular recognition. The "lock-and-key" theory of Fischer [189, 190] could be considered to be one of the earliest efforts to understand molecular recognition. Though the theory was originally developed to understand enzyme–substrate reactions, it was the basis of most of the initial attempts to explain molecular recognition. In this theory, the receptor–substrate complex results from a recognition of a substrate (key) by a molecular receptor (lock). However, with the advent of supramolecular chemistry and the concurrent efforts to develop molecule or atom specific hosts, it became evident that such a simplistic theory, which was purely based on steric considerations, was inadequate in explaining the affinity and specificity of hosts [191–193].

This understanding of the recognition phenomena becomes more acute in nanosystems wherein several competing interactions contribute to the recognition phenomena and hence contribute to the resulting structure and properties. Like in Figure 1, the formation of the organic nanotube is a result of the cooperative and competing interactions between two water molecules (hydrogen bonding) and between two aromatic molecules (π–π interactions) [16]. In general, the discussion of recognition phenomena is greatly facilitated, if we examine it in terms of contributions of individual molecular or atomic interactions. Though both intramolecular and intermolecular interactions influence recognition phenomena, the contribution of the latter is more significant. While a large number of interactions prevail in chemical and biological systems, the interactions which are relevant in the context of nanorecognition can be broadly classified into five different types. These include (i) hydrogen bonding, (ii) ionic interactions, (iii) intermolecular interactions involving π systems, (iv) metallic interactions, and (v) interactions involving quantum species. While it is difficult to segregate each of these interactions into water-tight compartments, we rationalize this classification in subsequent sections. It is, however, important to note that

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (367–404)

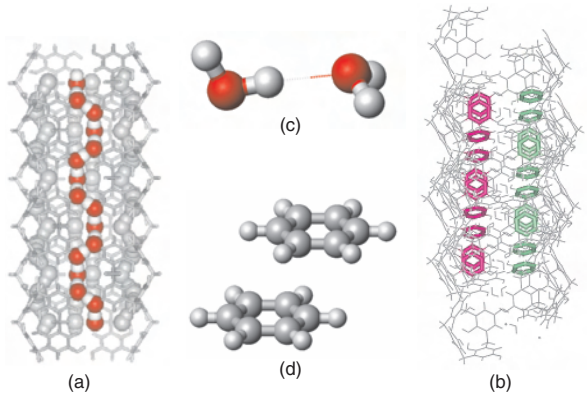


Figure 1. (a, b) A typical organic nanotube system consisting of water molecules and calix[4]hydroquinone is a result of juxtaposition of two competing intermolecular interactions: (c) H-bonding (water–water) and (d) π–π interaction.

recognition is a manifestation of several competing kinds of intermolecular interaction. This competition can either enhance or deplete the effect of each individual kind of interaction.

While it is often taken for granted, it should be pointed out that recognition per se does not imply binding/interaction [191]. It rather pertains to the structural, energetic, and thermodynamic events which precede the binding/interaction. Since the terms *binding* and *interaction* have been extensively used in literature, it is important at this juncture to distinguish the two terms. While *binding* can be said to result from the effects of several competing interactions, which could be of varying magnitudes, interaction is more specific and is either attractive or repulsive. Rather than attempting to discuss all the myriad aspects of nanorecognition, we focus our attention on the underlying physical principles governing the recognition phenomena.

Therefore, in the course of this chapter, we initially present a brief background on the development of the concept of recognition and we subsequently highlight the different aspects governing the recognition phenomena; (i) structural, (ii) energetic, and (iii) thermodynamic. The characteristics of different types of intermolecular interactions, which significantly influence recognition phenomena, are then discussed. Last, we elaborate on the competition between different kinds of intermolecular interactions and also discuss the concept of cooperativity.

Most experiments on nanosystems and for that matter on many host–guest complexes have been conducted in solid-state and solution. However, quantitative experimental investigations of most weak intermolecular interactions in complex systems are fraught with difficulties because of problems in isolating intermolecular complexes in well-defined orientations at room temperature [194–196]. In this connection, clusters, which are nothing but weakly bound aggregates of atoms, ions, or molecules, bridge the gap between atom and bulklike behavior, and hence results obtained from investigations of both atomic and molecular clusters are pertinent in understanding nearly all the principles of nanorecognition [197–233]. More compelling justifications are the numerous observations that data obtained on isolated clusters explain the bulk phase properties of nanomaterials [203] and also aid the design of new materials. In this connection, the recent design of novel nanomaterials and nanoelectromechanical devices based on results obtained from theoretical calculations of clusters merits mention [16–18, 233].

2. BACKGROUND

Contemporary interest in molecular recognition started after the pioneering discovery of the cation complexation abilities of crown ethers by Pedersen [234, 235]. Subsequent efforts by Dietrich et al. [236, 237], Gokel and Cram [238, 239], and others [7, 240, 241] revolutionized the field. Most of the early work on molecular recognition was predominantly devoted to the recognition of charged species like cations [240, 242–251]. This led to the discovery and development of a large number of host molecules (Fig. 2) capable of binding cations. The features of these hosts like podands [242], crown ethers [243–247], cryptands [248], spherands [249], and calixarenes [250] have been extensively discussed in the past. A number of hosts present in biological systems were also found to be excellent hosts for these charged species. Prominent among them are the siderophores [251], and molecules like valinomycin [252], enniatin [252], nonactin [252], etc. In most of these hosts, the cation binding is mediated through the heteroatoms (oxygen, nitrogen, sulfur) present in these hosts. While a number of practical utilities of these cation binding hosts exists [253, 254], an area of intense investigation has been the utility of these hosts in

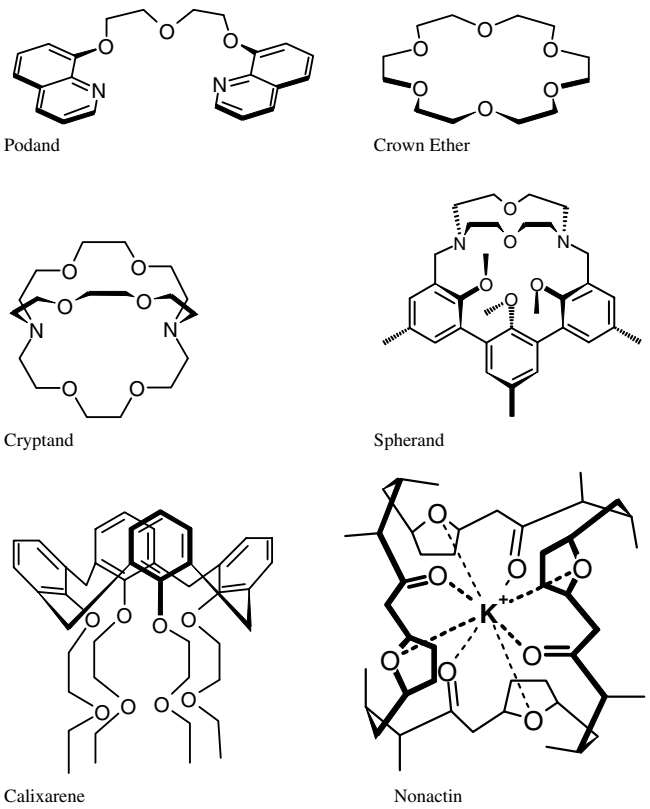


Figure 2. Some examples of cation binding hosts.

separating long-lived radioactive species from nuclear wastes [255]. This brings forth the twin issues of affinity and selectivity of these hosts toward a particular cation [56, 191, 196, 253, 254]. Most of the present work on these cation binding hosts is focused on developing environmentally benign and reusable systems, wherein the binding is reversible [253]. That is, the binding affinities can be influenced by the surrounding environment. In this context, a number of novel hosts wherein the binding of the cations to the hosts is mediated through olefinic or aromatic π systems present in the hosts have been developed. As an example, the [*n*]-beltenes which exhibit excellent affinity for the alkali metal cations are displayed in Figure 3. Apart from displaying excellent size selectivity (an issue which will be discussed in more detail in the next section), these hosts exhibit excellent affinity for a wide variety of cations and serve as useful models for understanding biological molecular recognition [256].

In the course of the development of hosts capable of binding cations, it was found that some of these hosts were also capable of binding neutral molecules and anions [241]. At the same time, a number of other host molecules like calixarenes [257], cyclophanes [258, 259], cyclodextrins [260, 261], cucurbiturils [75, 262], cryptophanes [196], and carcerands [263] were isolated, synthesized, and investigated for their neutral molecule or anion binding propensities. Interest in the binding of anions also stems from the fact that a number of biological processes involve molecular recognition of anionic species. In this context, the recent isolation of the X-ray structure of a chloride channel merits mention because it reveals the molecular basis of anion selectivity [264]. In the context of this chapter, it is interesting to elaborate briefly on the salient architectural differences between the Cl^- and K^+ channels [265–267].

The Cl^- channel has a roughly hourglass-shaped pore with a narrow constriction (its selectivity filter) near the membrane center (Fig. 4). On the other hand, the K^+ channel has

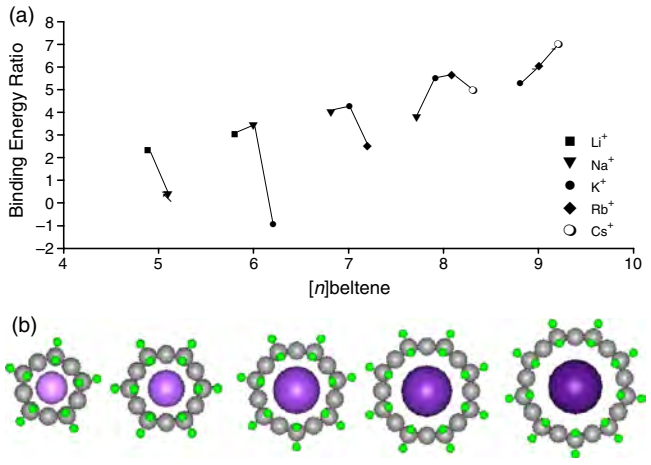


Figure 3. An example of a novel host (beltenes) binding cations through cation–π interactions. (a) The binding energy ratios of various [*n*]beltene-cation complexes. (b) Various cations captured by [*n*]beltenes (Li^+, Na^+, K^+, Rb^+, and Cs^+ are captured by [5], [6], [7], [8], and [9]beltene, respectively). Reprinted with permission from [256], H. S. Choi et al., *J. Org. Chem.* 67, 1848 (2002). © 2002, American Chemical Society.

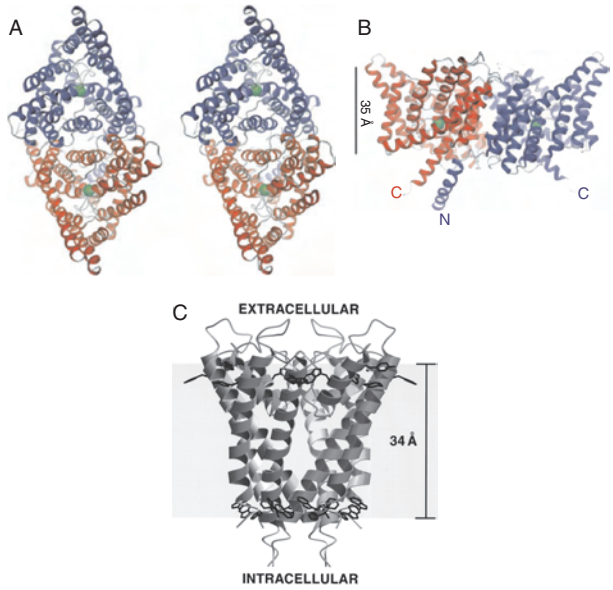


Figure 4. Architectural differences between the Cl^- and K^+ channel proteins. The X-ray structures of the potassium and chloride channels. Structure of the StClC dimer. (A) Stereo view of a ribbon representation of the StClC dimer from the extracellular side. The two subunits are blue and red. A Cl^- ion in the selectivity filter is represented as a green sphere. (B) View from within the membrane with the extracellular solution above. The channel is rotated by 90° about the *x*- and *y*-axes relative to (A). The black line (35 Å) indicates the approximate thickness of the membrane. (C) Ribbon representation of the potassium channel tetramer as an integral-membrane protein. Aromatic amino acids on the membrane-facing surface are displayed in black. Reprinted with permission from [264], D. A. Doyle et al., *Science* 280, 69 (1998), and [265], R. Dutzler et al., *Nature* 415, 287 (2002). © 1998/2002, American Association for the Advancement of Science, Macmillian Publishers Ltd.

a widening, or cavity, at the membrane center (Fig. 4). The cavity in K^+ channels, working in concert with helix dipoles, helps to overcome the dielectric barrier or energy barrier experienced by an ion owing to the low-dielectric membrane environment. The Cl^- channel, however, uses only helix dipoles. Though the origin of the interaction between the ion and the protein in both the K^+ and Cl^- channels is predominantly electrostatic, the difference in the geometry of the ion channels is due to the very different characteristics of the K^+ and Cl^- ions. We elaborate on this aspect in more detail in a subsequent section on ionic interactions.

Ion-pair recognition, the simultaneous complexation of cationic and anionic guest species by receptors, is a new and emergent field in contemporary nanorecognition. Thus receptors can be designed to exhibit novel cooperative and allosteric behavior, whereby the binding of one ion can have a profound influence on the coordination of the subsequent oppositely charged ion. These systems can serve as potential hosts for the selective extraction and transportation of environmentally important ion-pair species and also be used for zwitterionic recognition [268, 269]. In this context, it is interesting to highlight the recent design of amphiionophores (Fig. 5), which can bind both cations and anions [71, 72].

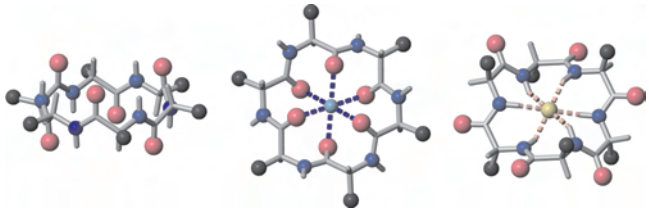


Figure 5. Example of an amphi-ionophore (cyclohexaalanyl) and its complexes with a cation (Na^+) and anion (F^-). Reprinted with permission from [71], S. B. Suh et al., *J. Phys. Chem. B* 106, 2061 (2002). © 2002, American Chemical Society.

While most of the previous discussion was focused on the recognition prevalent in single molecule hosts and guests, an area of equal importance is the macroscopic expression of molecular recognition (self-assembly of supramolecular and nanostructures) [191]. Essentially this implies that the focus is on the processes which follow the initial recognition event. These processes include (i) the formation of intricate shapes (templates), (ii) the recognition of these shapes or templates, and (iii) the template-directed self-synthesis of large polymolecular or nanoassemblies [66]. Therefore, molecular self-assembly can be defined as the spontaneous organization of molecules under thermodynamic equilibrium conditions into structurally well-defined and rather stable arrangements through a number of noncovalent interactions [191]. These molecules undergo self-association forming hierarchical structures. The key engineering principle for molecular self-assembly is to artfully design the molecular building blocks that are able spontaneously to undergo stepwise interactions and assemblies through the formation of numerous noncovalent weak chemical interactions (hydrogen bonding, ionic interactions, etc.). Although each of these interactions is rather weak, the collective interactions can result in very stable structures and materials. The key elements in molecular self-assembly are chemical complementarity and structural compatibility. This area has witnessed a lot of new developments in the recent past because of potential applications in the areas of liquid crystals, nanostructures of technological relevance, and the self-assembly, self-replicating characteristics of nano- and biostructures, nanoelectrical devices, and nanomechanical machinery [90–95].

Before we progress further, it is useful to highlight the distinction between the words supramolecule and supermolecule, because they are extensively used in the context of nanorecognition. Lehn has pointed out that the word "supramolecule" is used to describe large multiprotein architectures and organized molecular assemblies that result from the spontaneous association of a large undefined number of components into a specific phase having more or less well-defined microscopic organization and macroscopic characteristics depending on its nature (thin films, layers, membranes, vesicles, micelles, mesomorphic phases, solid state structures, etc.) [191]. The word "supermolecule" is used to denote singular or polymolecular association, which are well-defined, discrete oligomolecular species that result from the intermolecular association of a few components (receptor and its substrates bound together by molecular recognition) [191]. Therefore in the broad sense, supramolecular chemistry can encompass all types of molecular entitities ranging from well-defined supermolecules to extended, more or less organized, polymolecular associations.

3. PHYSICAL BASIS OF NANORECOGNITION

The principles governing nanorecognition are the same noncovalent interactions which occur in very weak complexes [232, 270], enzymatic and chemical transition states [136–148], and large molecular assemblies [92, 95]. The formation of nanostructures requires suitable specificities, optimal geometries, and reinforcing effects. Generally, nanorecognition is driven by enthalpic noncovalent forces (electrostatic, induction, dispersion) and by entropic effects [56]. A key role is played by the shape of the molecules or the molecular aggregates. While a few effects are highly specific and directional, others are less selective but are important factors in the stabilization of the eventual nanostructure. In self-assembly, a number of simultaneous interactions may occur, which would either enhance or deplete the binding at a particular site. Additionally cooperative and competitive effects play a vital role in determining the stabilitiy of the eventual nanostructure. It is important to stress that nanorecognition is a complex interplay of several structural, energetic, and thermodynamic components, each of which has a vital role in the stabilization of the observed structure.

3.1. Structural Considerations

The prevalence of the "lock and key" model around the time of the discovery of crown-ethers led to the notion that a close steric approach of the individual centers leads to optimal binding. This implied that structural/geometric considerations were enough to explain the enhanced ion selectivity of the crown-ethers. This together with the fact that the energetic effort needed by an open chain ligand to possess similar binding affinities as the corresponding macrocyclic ligand indicated that the preordained structural arrangement of the host might be an important factor governing the nature of binding [56]. However, it was soon realized that the presence of an enclosed cavity does *not* necessarily lead to increased stability even in cases where the solvent has no role to play in the interaction [56, 191]. The reason for this seemingly counterintuitive observation, which is based on the contribution of dispersive forces, is elaborated in subsequent sections. However, it is important to note that purely structural considerations though useful in explaining the binding affinities of some hosts is not very helpful in explaining most recognition phenomena. One has to take recourse to the role of both energetic and thermodynamic considerations.

3.2. Energetic Considerations

Recognition phenomena, as is observed in chemical and biochemical systems, is a juxtaposition of competing energetic elements which are either attractive or repulsive [56, 271]. Before we progress further, it is pertinent to note that a large number of terms have been employed to discuss various kinds of intermolecular forces [272],

prevalent in molecular recognition. In terms of relative importance, these intermolecular forces can be broadly described as (i) Coulombic forces, which describe the interaction between permanent charges, (ii) charge transfer forces, which describe the stabilization arising from a transfer of electrons from a highest occupied molecular orbital of one system to the lowest unoccupied molecular orbital of the other, (iii) dispersive or London attractions, which describe the attraction occurring due to the instantaneous distortions of the electron clouds of two interacting molecules. Though repulsive forces have contributions from several terms, the contribution from exchange–repulsion is the most dominant. Though not exactly equivalent, the electrostatic term describes the energies resulting from Coulombic forces, the induction term describes the energies resulting from polarization, charge redistribution, and charge transfer forces, the dispersion term describes the energies resulting from dispersive and London forces.

It is important to reiterate that the origin and magnitude of these individual attractive and repulsive elements is very different, and as will be shown in subsequent sections, are vital in governing the observed geometries and energies. Expectedly, the natures of the interacting species have important roles to play therein. Our knowledge of the nature and magnitudes of these individual interaction energies or intermolecular forces comes from two main sources: (i) quantum theory and (ii) experimental methods.

From quantum theory, we can compute, with varying levels of accuracy, what the potential energy is of a given configuration as a function of orientation, separation, etc. Quantum calculations enable us to determine the optimal geometry of the resulting species. There are a number of ways of performing quantum calculations, each with its own strength and weakness. Generally the more accurate, but computationally more intensive methods are those including electron correlation. Electron correlation energy is the energy recovered by fully allowing the electrons to avoid each other in a molecular ensemble. These correlation effects include both "dynamical" and "static" correlation effects. The inclusion of electron correlation is a prerequisite for evaluating the interaction energies because the magnitudes of both the induction and dispersion terms are critically dependent on it.

At the Hartree–Fock quantum level of theory, the interelectron repulsions are improperly treated in an averaged way [273]. As a consequence, it is of limited use in investigating the properties of most nanosystems. The methods like the second order Møller–Plesset level of theory, configuration interaction, and coupled cluster, which explicitly take into account the effect of electron correlation, are widely employed in investigating these nanosystems [232, 274, 275]. The only limiting factor, however, is the available computational resources.

Experimental methods also enable us to obtain the magnitude of individual interaction energy components. For example, dilute gas properties such as second virial coefficients, viscosities, etc. can be measured. Since most of these properties depend on the nature of the prevalent intermolecular interactions, the requisite information can be obtained by using standard equations along with a knowledge of the conditions under which the experiments were conducted. In this connection, the recent experimental determination of the magnitude of dispersion energies in a number of host–guest systems merits mention [194]. We can also obtain vital information from spectroscopy, molecular beam experiments, and more recently high resolution electron and atomic force microscopy [276–279]. The least elegant (but arguably most effective) method is to postulate a reasonable form of a forcefield, adjust the parameters for a simple system until a good fit with existing experiments is obtained, and then transfer this forcefield to other systems for use at other conditions [281].

It is useful to have an understanding of the origin of each of these intermolecular forces (electrostatic, induction, dispersion, exchange–repulsion) and a good starting point is standard perturbation theory [271, 282, 283]. The wave functions $|\psi^A\rangle$ and $|\psi^B\rangle$ of the isolated molecules A and B lead to a rigorous definition of the first, second, and higher order interaction energies in terms of the interaction Hamiltonian,

$$V^{AB} = \sum_{i\in A}\sum_{j\in B} q_i q_j / r_{ij} \tag{1}$$

In the equation, r_{ij} is the intermolecular distance. The electrostatic interaction between the two unperturbed molecules, which is equivalent to the first order term, is described by

$$\Delta E^{(1)}_{elec} = \langle \psi^A \psi^B | V^{AB} | \psi^A \psi^B \rangle \tag{2}$$

The second order terms, representing the induction and dispersion interactions, are obtained as an infinite sum over the excited states $|n^A\rangle$ and $|m^B\rangle$ of A and B:

$$\Delta E^{(2)}_{ind} = \sum_{n\neq 0} \left|\langle \psi^A \psi^B | V^{AB} | n^A \psi^B \rangle\right|^2 / (E^A_0 - E^A_n) + \sum_{m\neq 0} \left|\langle \psi^A \psi^B | V^{AB} | \psi^A m^B \rangle\right|^2 / (E^B_0 - E^B_m) \tag{3}$$

$$\Delta E^{(2)}_{disp} = \sum_{n,\,m\neq 0} \left|\langle \psi^A \psi^B | V^{AB} | n^A m^B \rangle\right|^2 / (E^A_0 - E^A_n + E^B_0 - E^B_m) \tag{4}$$

In these equations, E^A_i is the energy of the ith excited state of molecule A. In contrast to the equation for the dispersion energy, the two terms constituting the induction energy can be factored out because the integrations over $|\psi^A\rangle$ and $|\psi^B\rangle$ can be performed separately. Substituting the multipole expansion for V^{AB}, the induction energy is

$$\Delta E^{(2)}_{ind} = -1/2 \sum_{\alpha,\beta} \alpha^A_{\alpha,\beta} F^B_\alpha F^B_\beta - 1/2 \sum_{\alpha,\beta} \alpha^B_{\alpha,\beta} F^A_\alpha F^A_\beta - \text{higher order terms} \tag{5}$$

In this equation, the indices α and β denote the axes of the Cartesian coordinate frame. F^A_α is the electric field of molecule A at the multipole expansion center of the other molecule. The dipole polarizability $\alpha_{\alpha,\beta}$ describes the response to a fixed external field and is defined as

$$\alpha_{\alpha,\beta} = \sum_{n\neq 0} (\langle \psi | \mu_\alpha | n \rangle \langle n | \mu_\beta | \psi \rangle + \langle \psi | \mu_\beta | n \rangle \langle n | \mu_\alpha | \psi \rangle) / (E_n - E_0) \tag{6}$$

where μ_α are the components of the electric dipole operator. While both the electrostatic ($\Delta E^{(1)}_{\text{elec}}$) and induction ($\Delta E^{(2)}_{\text{ind}}$) energies are obtained in the fixed potential of each constituent molecule, the dispersion energy ($\Delta E^{(2)}_{\text{disp}}$) is closely associated to the coupled fluctuations of the charge distributions of both the constituent molecules A and B. In contrast to these attractive energies, the exchange energy which is repulsive for closed shell complexes emerges from the fermionic nature of electrons. Since Fermi–Dirac statistics requires antisymmetrization of the product functions $|n^A\rangle|m^B\rangle$, the first order exchange energy is usually defined as

$$\Delta^{(1)}_{\text{exch}} = \frac{\langle \psi^A\psi^B|SV^{AB}|\psi^A\psi^B\rangle}{\langle \psi^A\psi^B|S|\psi^A\psi^B\rangle} \tag{7}$$

or

$$\Delta^{(1)}_{\text{exch}} = \frac{\langle S\psi^A\psi^B|H^0 + V^{AB}|S\psi^A\psi^B\rangle}{\langle S\psi^A\psi^B|S\psi^A\psi^B\rangle} - E^A - E^B \tag{8}$$

where H^0 is the unperturbed Hamiltonian and E^A and E^B are the ground state energies of A and B. When both $|n^A\rangle$ and $|m^B\rangle$ are exact eigenfunctions, these equations are equivalent. There is no unique term for higher order terms of the exchange energy. This is because the antisymmetrizer S mixes different orders of perturbation theory as S and V^{AB} do not commute. Moreover, the projected functions $S|n^A\rangle|m^B\rangle$ generally form a linearly dependent set. The available evidence indicates that second order exchange effects (correction of the induction and dispersion energies due to exchange) are small for most intermolecular complexes. This implies that the addition of long range energies to the exchange interactions does not lead to serious errors. In multimolecular assemblies, the total interaction energy can be written as a sum of the pairwise interaction energies (pairwise additivity). While the first order electrostatic energy is exactly additive, the dispersion energy is nearly additive (the three-body Axilrod–Teller term is generally small in the third order). In contrast, the induction energy is not pairwise additive because of interference of electrostatic fields from different sources. Furthermore, the higher order contributions in complexes containing polar molecules cannot be neglected because the induced fields can significantly alter the molecular wave functions. This in turn affects the additivity of exchange interactions because its magnitude is related to the extent of intermolecular integral overlap.

3.3. Thermodynamic Considerations

While the origin of these intermolecular forces can be understood using quantum theory, one needs to use formulas derived from statistical thermodynamics to compute them as functions of temperature. This is because most calculations based on quantum theory invoke the approximation that the atomic nuclei are stationary, with the electrons swarming about them. This is a good approximation because nuclei are much heavier than electrons. However, the resulting energies are for a hypothetical, nonvibrating molecule. Though these oscillators can be considered to be at rest in a classical mechanics model, the real quantum-mechanical oscillators are always in motion. The small residual motion at absolute zero temperature is the zero-point vibrational energy.

One can obtain the macroscopic thermodynamic properties like enthalpies, entropies, and heat capacities using the molecular partition function. The molecular energy levels ϵ_i are used to compute the molecular partition function, usually denoted by the symbol Q, as shown in the following equation:

$$Q(T) = \sum_i \exp(-\epsilon_i/kT) \tag{9}$$

The sum extends over all energy levels. However, for very high temperatures at which the molecule or molecular complex becomes unstable, the extent of the sum may be ambiguous. One should therefore note that most theoretical estimates of thermodynamic properties (i) depend strongly upon the high-energy cutoff procedure adopted and (ii) deviate implicitly from the ideal-gas model. One typically chooses the lowest energy level to be the zero of energy, so that no levels lie at negative energies. From Eq. (9), it follows that the largest contributions to Q are from the lowest energy levels. Conversely, levels that lie far above kT (207 cm^{-1} at room temperature) have only a minor effect on Q and its derivative thermodynamic quantities:

$$S = Nk\left[\frac{\partial}{\partial T}(T\ \ln Q) - \ln N + 1\right] \tag{10}$$

$$C_v = NkT\frac{\partial^2}{\partial T^2}(T\ \ln Q) \tag{11}$$

$$C_p = C_v + R \tag{12}$$

$$H(T) - H(0) = \int_0^T C_p\, dT = NkT^2\frac{\partial(\ln Q)}{\partial T}_{N,V} + NkT \tag{13}$$

Given the molecular partition function, the usual molar thermodynamic functions can be calculated based upon the following general equations. Equation (10) is for the entropy, (11) for the heat capacity at constant volume, (12) for the heat capacity at constant pressure, and (13) for the enthalpy difference relative to absolute zero temperature. N is Avogadro's number, k is the Boltzmann constant, and the ideal-gas constant $R = Nk$. The last two terms inside the brackets in Eq. (10) arise from the indistinguishability of identical molecules, which requires a factor of $(1/N!)$ in the partition function for the ensemble.

The effect of temperature on intermolecular forces can be explained using the concept of thermal energies (kinetic theory). Thus, the intrinsic stability of an intermolecular complex at a given temperature is due to the balance between their thermal and potential energies. It is of interest to note that routine physical phenomena like melting, boiling, and the presence of critical points reflect the magnitude of the intermolecular interaction. For example, the critical temperatures of neon dimer (44 K), methane dimer (191 K), and water dimer (647 K) mirror their binding strengths [280]. Similar trends can also be noted for the molar heat of vaporization.

Though energetic considerations enable the formation of the intermolecular or supramolecular complex, thermodynamic factors play a crucial role in determining the stability of the complex. As was mentioned earlier, the individual monomers involved in the recognition process undergo substantial conformational changes upon binding. These conformational changes result from the monomers trying to maximize the intermolecular interaction and minimize the induced intramolecular conformational strain. In cases where solvent is involved, desolvation effects also have an important role. All the aforementioned effects influence the nature of the intermolecular interaction through the entropy term. Experimentally, the major thermodynamic quantities like the free energy (ΔG) and enthalpy (ΔH) are obtained either by calorimetric titration or from the temperature variation of the complex stability constant (K). The complex stability constants (K) at different temperatures are obtained by using any experimental technique like potentiometry, conductometry, spectrophotometry, nuclear magnetic resonance (NMR) spectrometry, solvent extraction, etc. Based on a plot of the logarithm of the stability constant (K) and the inverse temperature ($1/T$), the enthalpy and entropy changes are obtained as the slope and intercept, respectively.

$$\ln K = \frac{\Delta H}{NkT} + \frac{\Delta S}{Nk} \tag{14}$$

Based on experimental information obtained from a large number of intermolecular complexes, it emerges that entropic effects are much more dominant in complexes predominated by electrostatic interactions. We discuss the theoretical basis of this observation in subsequent sections.

4. VARIOUS TYPES OF INTERMOLECULAR INTERACTIONS

4.1. Hydrogen Bonding

4.1.1. Definition, Importance, and Classification

Hydrogen bonding, which is one of the most widely prevalent and extensively investigated intermolecular interactions, is typically characterized by an equilibrium configuration involving a hydrogen atom donor **D** and an acceptor **A** [284–287]. Conventionally, the hydrogen bond (H-bond) is represented as (**D**–**H**$\cdots$**A**), wherein both donor **D** and acceptor **A** are electronegative atoms. One can broadly classify the experimentally observed H-bonds into three classes (weak, medium, and strong) based on the strength of their interaction. This strength, based on experimental and theoretical methods, ranges between 0 and 30 kcal/mol ($\sim$30 kcal/mol in case of the $F^- \cdots H$–F interaction). In general, weak H-bonds involve polar X–H groups like C–H or P–H, or less polar acceptors like covalently bonded fluorine atoms or π systems [232, 288, 289]. The vast majority of H-bonds discussed in literature are of medium or intermediate strength ($\sim$5 kcal/mol). Thus, the H-bonds observed in water clusters, carboxylic acid dimers, H-bonds involving uncharged species in proteins, and nucleic acids are well-known examples [290–293]. Strong H-bonds in general involve a charged species like an oxyanion as a proton acceptor and are mostly observed in enzyme active sites [294]. Though the high interaction energies involved therein have elicited a lot of debate on the exact nature of H-bonds, the properties of systems exhibiting these H-bonds have attracted a lot of interest in the context of molecular nanowires, switches, etc. Since the interaction in these strong H-bonds is predominantly mediated through electrostatic interactions, we discuss their characteristics in the section on ionic interactions.

The major reasons for the importance of H-bonding in nanorecognition are that (i) the interaction is of intermediate strength and therefore reversible, (ii) the interaction is directional and therefore one-, two-, three-dimensional structures can readily be assembled by hydrogen bonding, and (iii) the assembly is often fast and specific. As the interaction is weaker than a covalent bond, the bonding can be reversed and hence error correction can be carried out. The omnipresence of H-bonds in most biological molecules, viz. proteins and nucleic acids, is due to the aforementioned characteristics. However, the stellar role of H-bonding can be observed in the properties of water and ice.

4.1.2. Origin and Characteristic Features

Though the first proposal of H-bonding was put forward by Moore and Winmill [295], the concept gained widespread attention and recognition after the work of Latimer and Rodebush [296]. However, Pauling was the first to explain the concept of H-bonds using the example of water [297]. Thus, in comparison to similar small molecules, water might be expected to be a gas under everyday conditions—which would make the world a very different place. But water remains a liquid because H-bonds impart to the molecules extra cohesion, preventing them from flying apart into steam. In the early 1930s, Pauling proposed that these H-bonds were the result of an uneven distribution of electrical charge in the water molecule (H_2O). This was attributed to the greater electronegativity of oxygen, as a result of which the electrons cling more closely to the oxygen atom than to the hydrogen atoms. This leads to the oxygen atom being more negatively charged than the hydrogen atoms. As a result of electrostatic interactions, the hydrogen atoms of adjacent water molecules tend to stick to the oxygen atom of the neighbor. Pauling subsequently tried to explain some of the anomalies of the electrostatic theory by proposing that a small degree of electron sharing between the two water molecules contributes to the H-bond [298]. It should be noted that sharing of electrons between atoms generally results in strong, so-called "covalent bonds," like those that hold the two hydrogen atoms to an oxygen atom in individual water molecules. In the intervening years, since the first explanations of the H-bond were given by Pauling, a large body of information gleaned from both experimental and theoretical studies has led to a detailed understanding of the nature of the H-bond [284–287]. Most of these studies have incidentally been carried out on the simplest intermolecular complex containing a H-bond (water dimer) or larger clusters of the water molecules [299–323].

Incidentally, it has also been shown in the course of these investigations that most of the properties of water in the

condensed phase can be obtained from the properties of these individual water clusters. In Figure 6, the magnitudes of the different interaction energy components at different intermolecular separations of the water dimer are shown. It can be seen that the contribution of electrostatic energy is the largest attractive contributor (60%) to the interaction energy. Expectedly, the other attractive contributions (dispersion and induction) are of the same order of magnitude as the total interaction energy [321]. This implies that the magnitude of the repulsive contribution is nearly as high as the total electrostatic energy. In fact, it has been noted that the global minimum of the water dimer and similar H-bonded structures corresponds to a structure with the largest attractive contributions and a large exchange–repulsion term [324]. The local minima and the transition states, in turn, show smaller attractive terms and lower repulsion with respect to their values for the global minimum. Thus the decrease of exchange–repulsion is as important as the increase of the attractive terms.

Most of the characteristic properties of the H-bond, which are observed experimentally, can be explained to a large extent by this physical model. These characteristic properties, which provide the means to identify the presence of H-bonds, include changes in the bond lengths, dipole moments, vibrational frequencies, etc.

The **D**–H bond length is often increased upon the formation of a H-bond. These increases are particularly significant in crystal structures exhibiting H-bonds. The fundamental frequency associated with the **D**–H bond stretch is often lowered upon formation of a complex. This redshifting is one of the most sensitive probes of H-bonding in experimental systems. The origin of this effect can be attributed to the fact that the fundamental frequency of the **D**–H vibration in an isolated monomer is proportional to the energy difference between the ground ($\upsilon = 0$) and the first excited ($\upsilon = 1$) states. However, in the complex, the frequency would be different because the intermolecular van der Waals vibrations

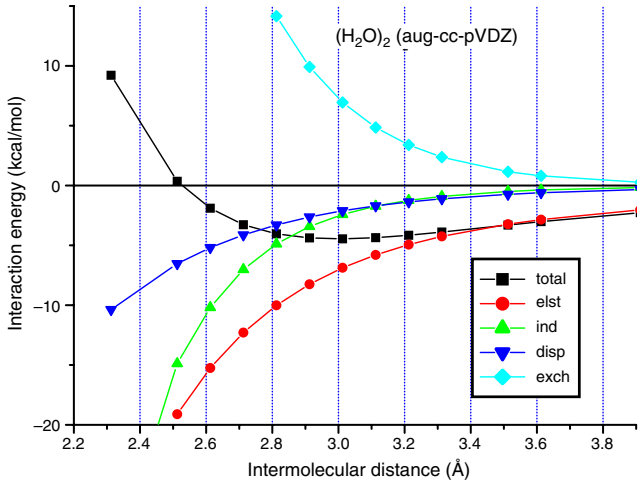


Figure 6. Magnitude of the interaction energy components of the water dimer at different intermolecular separations. The calculations have been carried out using the aug-cc-pVDZ basis set (elst: electrostatic energy, ind: induction energy, disp: dispersion energy, exch: exchange energy).

also have to be taken into account to evaluate the energies of both the ground and first excited states. The first excited state potential is always deeper in a H-bonded system because of an increase in the magnitude of the electrostatic and induction energies. Recent calculations, however, indicate the importance of dispersion energies in the magnification of the excited state potential of the H-bonded complex [325]. The H···**A** distances in H-bonded systems are very short, compared to other intermolecular bonds. This is a consequence of the strength of the H-bonds and incidentally also helped develop the concept of the H-bond. Thus, the hydrogen bonds are much shorter than the van der Waals radii of hydrogen and atom **A**. These van der Waals radii were deduced by Pauling, using the distances between the atoms in molecular crystals as a guideline [298].

It has been shown by many authors that nearly all known variants of the H-bond with widely different magnitudes can be explained using this model. Thus, the C–H···O interactions are very similar to conventional O–H···O H-bonds in most respects, but are generally weaker [87, 326]. A hydrocarbon like methane exhibits a very weak interaction with a proton acceptor such as water, with the interaction energy being only a fraction of a kilocalorie per mol. However, what distinguishes the O–H···O and C–H···O interaction is the intermolecular separation. In the former, the intermolecular O···O distance is of the order of 2.9 Å, but in the latter, the intermolecular C···O distance is of the order of 3.5 Å. Interestingly, it was also shown that electronegative substitutents on the donor carbon, like fluorine, add about 1 kcal/mol to the interaction energy and draw the two subunits together by 0.10–0.15 Å. Furthermore, the strength of the interaction is found to be much less sensitive to the nature of the proton acceptor than to the donor. Nearly all kinds of H-bonds exhibit very similar characteristics as regards their structural features. Thus, both the O–H···O and C–H···O H-bonds are disposed to a linear arrangement of the three atoms. However, the latter can be easily bent.

In the context of the redshift of the **D**–H vibrational frequency shifts upon H-bonding, mention should be made of some systems like the C–H···O interactions which exhibit a characteristic blueshift [327–329]. This blueshifting of the C–H stretch led to the characterization of the C–H···O H-bond as an "anti H-bond" [329]. However, recent studies on several systems exhibiting such a blueshifting of the **D**–H frequency shifts indicate that this blueshift is a consequence of the differing magnitudes of the individual interaction energy terms [326]. The change in **D**–H bond length is the net result of one set of forces tending toward elongation and another that pulls toward a shorter bond. While the attractive forces (electrostatic, induction, and dispersion energies) tend to push the hydrogen away from the donor atom, exchange–repulsion pulls it away from the acceptor. In the case of the O–H···O interaction, the attractive forces are slightly stronger than the exchange–repulsion, but in the case of the C–H···O H-bond, the exchange–repulsion is nearly similar in magnitude to the attractive forces or very narrowly overcomes them. The blueshift in the C–H···O is therefore a consequence of this bond contraction [326].

In the following section, we highlight some of the interesting features of H-bonds using the example of water clusters. The smallest water cluster, the water dimer, possesses a

linear structure. However, as the size of the water cluster increases, one notes interesting changes in their structures. Thus, the water trimer, tetramer, and the pentamer form cyclic structures. The hexamer heralds the onset of three-dimensional structures, a trend which extends to the larger water clusters (Fig. 7) [226, 315].

As can be seen from Figure 7, an increase in the size of the water cluster leads to an increase in the number of H-bonds. However, this increase in the number of H-bonds need not necessarily translate itself into enhanced stability. The orientation and the nature of the H-bonds also play an important role in governing the stability of the H-bonded structure. A case in point is the water pentamer. Thus, the pyramidal structure *R444*, which has six H-bonds, is nearly 3.1 kcal/mol higher in energy than the cyclic *R5* structure (Fig. 8). Furthermore, the *R33* structure, which has the same number of H-bonds as the *R444* structure, is nearly 3.3 kcal/mol higher in energy. The enhanced stability of the *R5* structure can be attributed to the role of nonadditive effects. The importance of this nonadditive effects in the energies and hence the geometries of the water clusters emerges from the earlier investigations of one of the authors

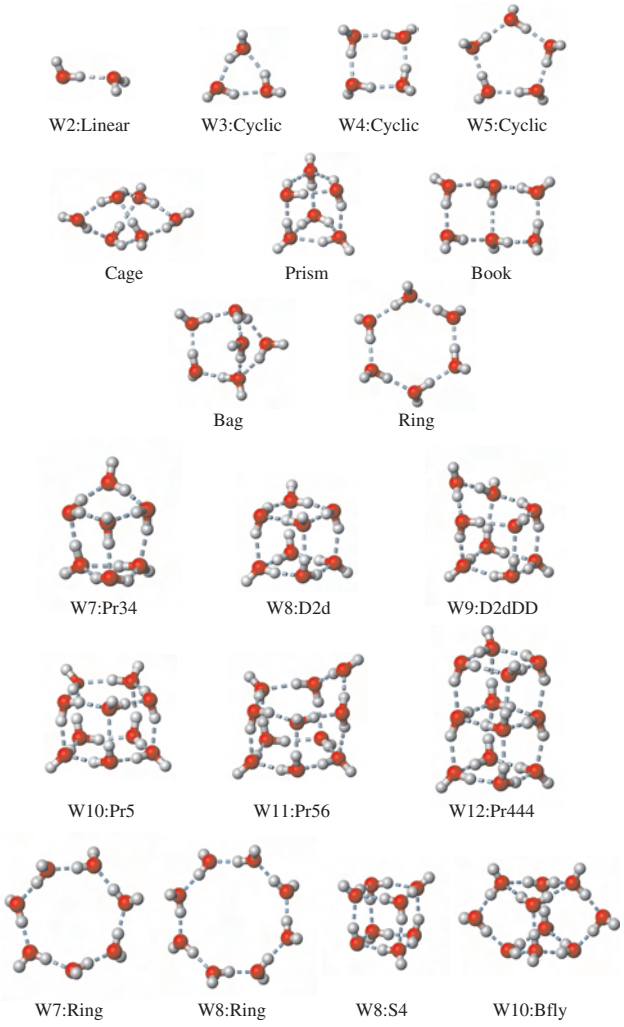


Figure 7. Structure of the water clusters from dimer to dodecamer. Note that the hexamer heralds the onset of 3D structures.

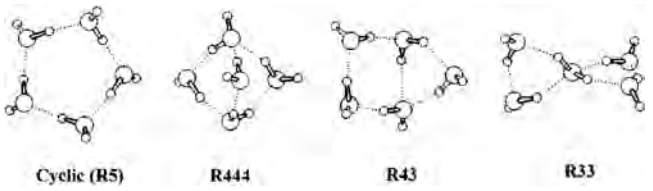


Figure 8. Various structures of the water pentamer. Presence of more H-bonds does not result in higher stability. The stability is governed by the competition between H-bond and strain energy.

[228], which led to the successful prediction of the lowest energy conformers of the water hexamer, heptamer, and octamer.

An important consequence of the presence of multiple H-bonds is cooperativity [303, 318, 330]. While its role in nanomaterial design is elaborated subsequently, we highlight some of the effects of cooperativity here. One of the most important consequences of cooperative interactions is the enhancement of the H-bond strength. Thus, in case of the water clusters, the average energy of all cluster H-bonds increases significantly (Fig. 9) with cluster size (trimer 3%, tetramer 36%, pentamer 44%, and hexamer 49%). The cooperative enhancement of the interaction energies also leads to significant geometry changes, with a progressive decrease in the intermolecular H-bonding distance. One can see this in more detail in Figure 10, wherein we have plotted the average intermolecular distance R_{O-O} as a function of the cluster size. Note that beyond a certain cluster size, there is no further decrease in R_{O-O}. On a similar note, the proton NMR shifts of these water clusters are shifted downfield as function of cluster size. In Figure 9, one can also note the consequences of cooperative effects on both the enthalpies and free energies. An interesting observation in the context of cooperative effects in water clusters is that most of the properties are nearly saturated at the water hexamer. This observation can be explained by the fact that all structures beyond the water hexamer are three-dimensional, and the OH frequency shifts are saturated at the hexamer [226, 315].

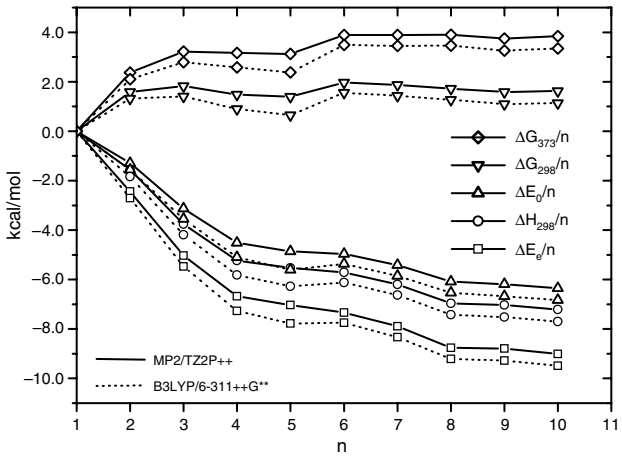


Figure 9. Role of cooperative interactions on energetic stabilization of larger water clusters. (ΔG: free energy, ΔH: enthalpy, ΔE: internal energy. The subscripts correspond to the temperature in Kelvin.)

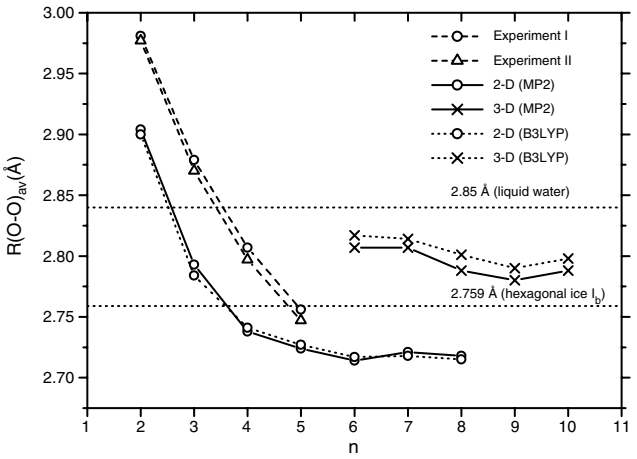


Figure 10. Geometric effects of cooperative interactions. An examination of the R_{O-O} distance as a function of the size of water clusters. The calculations have been carried out at the MP2/TZ2P++ and B3LYP/6-311++G** levels of theory [226].

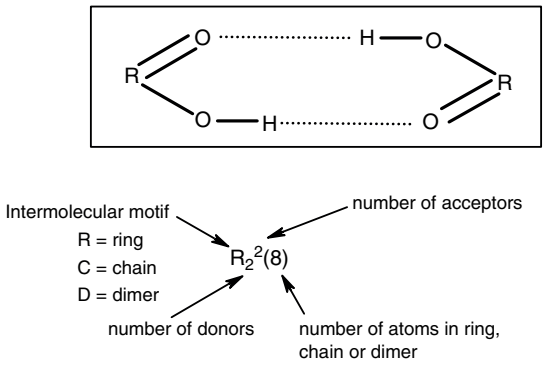


Figure 11. Components of a H-bonding graph set.

While most of the discussion on H-bonding was focused on the origin and their influence on various properties, it is useful to discuss their role in molecular assemblies. In this connection, the H-bond classification approach of Etter et al. is particularly valuable because it relies on the intermolecular specificity of H-bonds and deliberately ignores competing factors like preorganized cavities, ionic interactions, and steric interactions [331]. Based on graph theoretical concepts, they describe complex networks involving many different types of H-bonds into simpler motifs. Interestingly, the myriad patterns of H-bonds evolve from just four simple motifs. A notation describing these motifs is assigned to each different type and provides a complete and accurate description of the patterns of H-bonds in a molecular assembly at any desired level of complexity [331].

The process of assigning a graph set begins from an identification of different types of H-bonds contained in a given network. A *motif* is defined as a pattern formed by connecting molecules by just one type of H-bond. The *motif* is further characterized by one of four *designators*. These designators are C (chain), R (ring), and D (dimer) for motifs generated from intermolecular H-bonds. The S (self) designator denotes an intramolecular H-bond. The number of donors and acceptors are assigned as subscripts and superscripts, respectively. The motif size (denoted as the *degree*) is determined by the number of atoms in a ring or the repeat unit of a chain and is given in parantheses following the designator of the pattern. The procedure for assigning the graph set of an eight-membered ring formed by carboxylic acids is shown in Figure 11. In Figure 12, various motifs along with their graph sets are shown.

Based on the analysis of the graph sets of a large number of crystal structures, it soon became evident that H-bonds in organic molecules containing functional groups follow three general rules. They are: (i) all good proton donors and acceptors are used in H-bonding; (ii) intramolecular H-bonds that form a six-membered ring are preferred over intermolecular H-bonds; and (iii) the best proton donors and acceptors remaining after formation of intramolecular H-bonds form intermolecular H-bonds to one another. A number of additional rules pertaining to special classes of functional groups are also known [57]. The discussion in the preceding paragraphs clearly reveals that these rules are

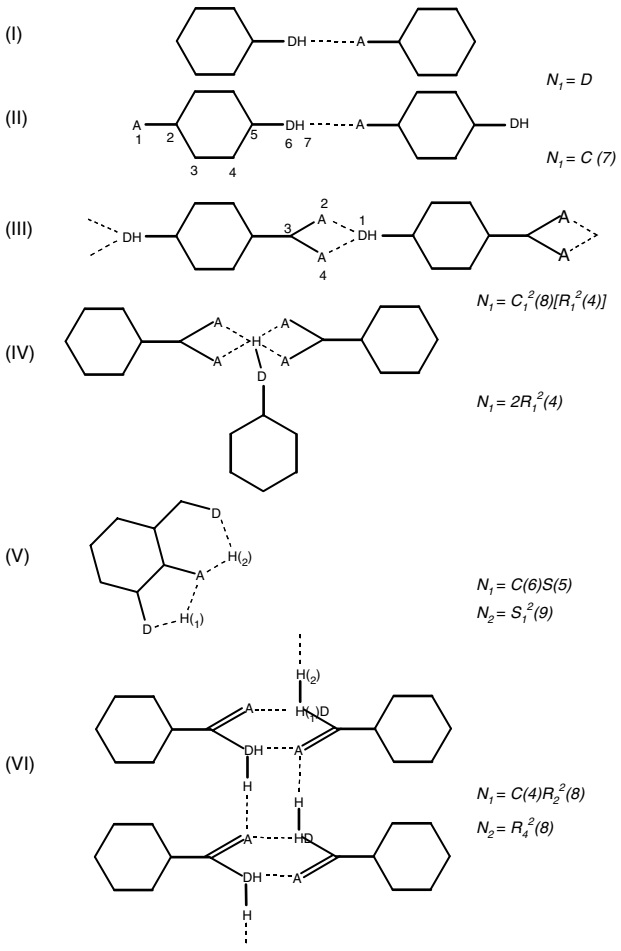


Figure 12. Examples of different motifs of H-bonds with their graph sets. The hexagons are meant to represent any organic ligand. This specific form is given so that the degrees of the patterns can be assigned in this figure. If the organic ligands were other kinds of groups, then the degrees of some of the patterns would change. Reprinted with permission from [331], M. C. Etter et al., *Acta Cryst. B* 46, 256 (1990). © 1990, International Union of Crystallography.

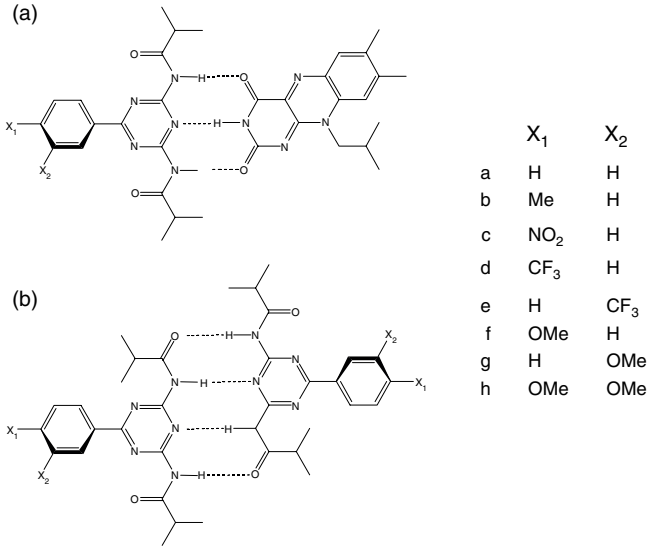
based purely on energetic considerations. A number of other factors which might influence H-bonding are steric effects, competing ionic forces, and the presence of multiple sites for H-bonding. The latter two are discussed in more detail in subsequent sections.

4.1.3. Manifestation

Given these characteristics of the H-bond, it is of interest to examine how these interactions have been employed in nanomaterial design. While several recent reviews exist on this subject, we highlight a few strategies employed to enhance the efficiency and selectivity of H-bonded interactions.

As was mentioned earlier, of all the three attractive interactions, electrostatic interactions play a major role in the stability of the H-bond. Therefore, a strategy, already discussed, would be to create a system wherein the positive charge upon the donor proton is increased and/or the negative charge upon the acceptor atom is increased. Additionally, the strong dependence of the H-bond on electron distribution implies that substitutent effects provide an effective means of modulating the strength of the interaction.

Indeed, this was the strategy pioneered by Wilcox and co-workers [332] and more recently studied in more detail in a complex system by Rotello and co-workers (Fig. 13) [333, 334]. Thus, the authors observed that the recognition of guests (Fig. 13a) and the dimerization of hosts (Fig. 13b) were found to be governed by the nature of the substituents. Thus the dimerization constants (K_{dim}) measured in $CDCl_3$ range from 95 M^{-1} for the host with a phenyl substituent to 41 M^{-1} for the *p*-nitrophenyl-substituted host. Interesting was the observation that the electron donating groups on the host enhanced dimerization and electron withdrawing substituents diminished recognition over a rather narrow (0.5 kcal/mol) range. On the other hand, the affinity of the guest was enhanced with electron withdrawing groups and decreased with electron donating substituents. This relationship between substituent and guest affinity in this above host–guest system can be understood if one examines the binding surface of the host. An increase in the electron density at the central nitrogen of the **DAD** surface of the host enhances H-bonding at this site and was experimentally confirmed by the observance of the maximum shifts for the N(3) proton of the guest (increases from 12.2 ppm for **b** to 19.8 ppm for **h**). This increase in the electron density, however, lessens the positive charge on the donor amide protons and hence diminishes the interactions at these two positions. In the course of these studies, it can also be seen how molecular recognition or H-bonding processes provide a means of investigating electronic communication in complex systems, which incidentally is a prerequisite for molecular wires and other devices.



	X_1	X_2
a	H	H
b	Me	H
c	NO_2	H
d	CF_3	H
e	H	CF_3
f	OMe	H
g	H	OMe
h	OMe	OMe

Figure 13. Examples of modulation of H-bond strengths through substituents. (a) Host–guest binding through a **DAD–ADA** binding surface. (b) Host dimerization through **ADAD–DADA** binding. Reprinted with permission from [334], G. Cooke et al., *Chem. Soc. Rev.* 31, 275 (2002). © 2002, Royal Society of Chemistry.

The cumulative free energy of a H-bonded complex is the difference between energy required by the host or the guest to reorganize themselves to achieve a complementary binding site with each other and the enthalpically favorable interaction energy upon host–guest complexation. Thus, when the donor and the acceptor are preorganized, there is a dramatic reduction in the overall free energy of complexation because the energy required for the hosts or guest to conformationally reorganize themselves is minimized.

Though there have been a number of reports of the development of conformationally preorganized hosts [335–349], the work of Meijer and co-workers merits mention because they have spectacularly used the cues from DNA base pairing to develop controlled molecular architectures [336–345]. Basically the work involves the development of host–guest systems that form arrays of donor and acceptor atoms which can undergo spontaneous homo- and heterodimerization through the formation of quadruple H-bonds. Incidentally, quadruple H-bond arrays have been predicted to be much stronger than triple H-bond arrays [346]. These investigations also include efforts to minimize any possible repulsive interactions between the donor–donor (**D**$\cdots$**D**) and acceptor–acceptor (**A**$\cdots$**A**) sites.

Of all possible motifs which resulted from homo- and heterodimerization, the **DDAA** arrangement was found to form more effective H-bonding complexes than the **DADA** arrangement. This is because of the fact that the latter has only six repulsive secondary interactions in contrast to the two repulsive and six attractive secondary interactions of the former (Fig. 14). In the ureidopyrimidone systems, the pyridimin-4(1H)-one tautomerizes to the pyrimidin-4-ol form leading to the formation of dimers via preorganized **DDAA** or **DADA** arrays (Fig. 15), respectively. In both cases, the preorganization through the formation of an intramolecular H-bond results in the formation of self-assembled systems by linear arrays of four H-bonds. The K_{dim} in these preorganized systems are much larger than those predicted for non-preorganized arrays [343]. The juxtaposition of both preorganization and minimization of repulsive secondary interactions leads to the **DADA** array having a K_{dim} of 10^7 M^{-1}, which is nearly 100 times larger than the observed K_{dim} of 10^5 M^{-1} in the **DADA** array [343]. Meijer et al. also noted that the H-bonding propensities of

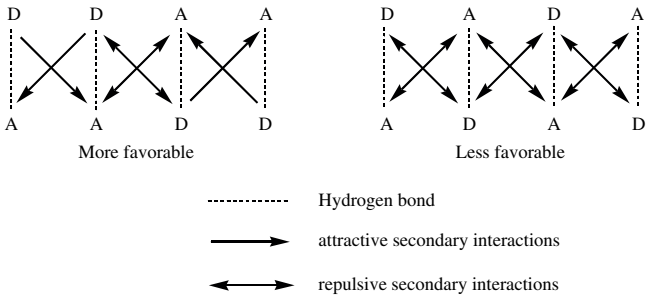


Figure 14. Attractive and repulsive secondary interactions in quadruple hydrogen bonded systems. Reprinted with permission from [334], G. Cooke et al., *Chem. Soc. Rev.* 31, 275 (2002). © 2002, Royal Society of Chemistry.

these ureidopyrimidone systems could be further tweaked by changing the electronic characteristics of the functionality in the sixth position of these heterocycles. Thus, electron donating groups result in the formation of complementary **DDAA** arrays characteristic of the pyridimin-4(1H)-one form. On the other hand, electron withdrawing substituents lead to the formation of **DADA** arrays, which is characteristic of the pyrimidin-4-ol form. In extensions of this work, helical self-assembled polymers were made by harnessing the quadruple H-bonding propensities of the ureidotriazines [343].

In the context of preorganization directed H-bonding, mention should be made of a recent work wherein the preorganized structure of calix[4]hydroquinone was used to design both one-dimensional and two-dimensional H-bond arrays [16, 18]. Interestingly this design strategy also led to the revealation that the electrochemical and photochemical properties of the parent calix[4]quinone molecule could be harnessed to design novel redox modulated H-bonds.

The strategy of redox modulation of the H-bonding process has, however, been employed in a number of previous instances [61, 350–354]. Redox processes are widely prevalent in biological systems with a number of crucial enzymatic reactions being mediated through changes in the redox states of enzyme cofactors (quinones, flavins, nicotinamides, etc.) [350, 353]. In an effort to understand the mechanism of the flavin enzymes, Breinlinger and co-workers examined the interaction of flavin with a 2,6-diaminopyridine derivative under different redox conditions (Fig. 16) [355]. What they noted was that a reduction of flavin to the corresponding flavin radical anion led to a nearly 500-fold increase in the binding constant. The role of the central N–H···N H-bond in affecting the binding constants was discounted by carrying out the binding studies in a range of H-bonding solvents, thereby confirming that the enhanced binding resulted entirely from changes in the H-bonds involving the carbonyl oxygens [355]. The use of redox modulation of H-bonds in aiding the development of addressable surfaces with device applications was also demonstrated in a recent study [354].

There have been a number of recent reports on the use of photochemical means to influence the H-bonding formation [355–357]. However, most of these changes were brought by harnessing the excited-state characteristics of π systems. Therefore, we would discuss these systems in subsequent sections devoted to intermolecular interactions involving π systems.

In the course of this section, we have examined the origin, characteristics, and different manifestations of H-bonding in various systems. In the next section, we elaborate on how the characteristics of H-bonding can be modulated by interaction with anions and cations.

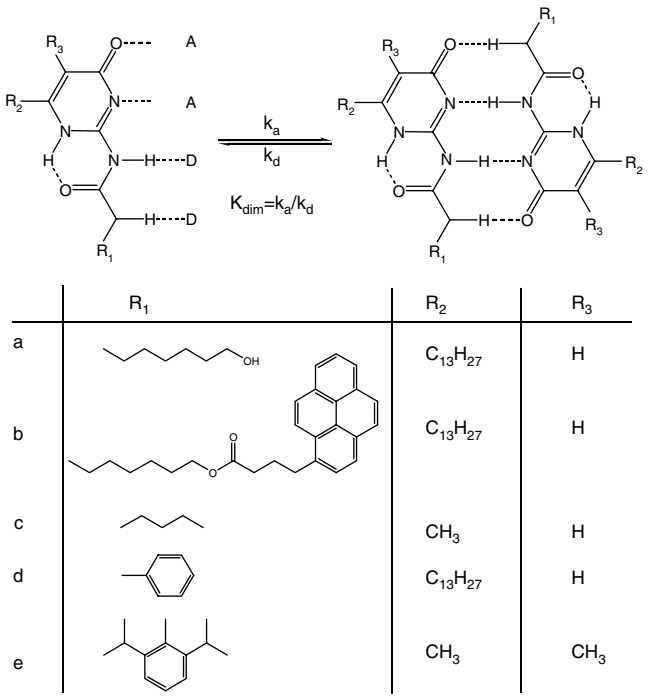


Figure 15. Dimerization of 2-Ureido-4[1H]-pyrimidinone. Reprinted with permission from [343], S. H. M. Söntjens et al., *J. Am. Chem. Soc.* 122, 7487 (2000). © 2000, American Chemical Society.

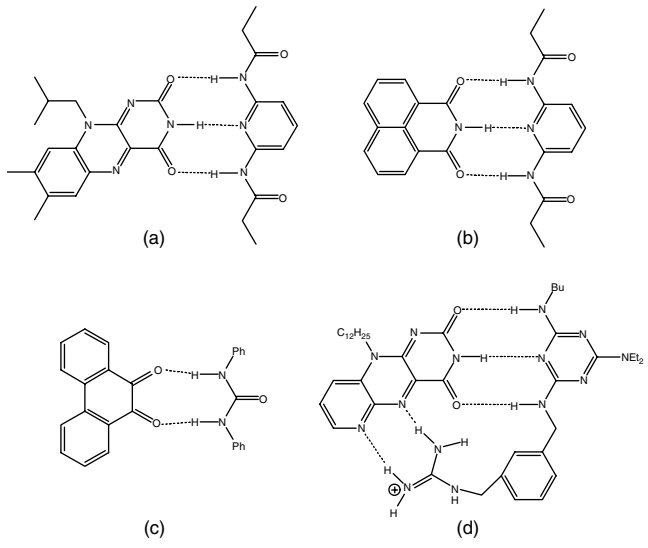


Figure 16. Interaction of flavin with 2,6-diaminopyridine. Reprinted with permission from [355], E. Breinlinger et al., *J. Am. Chem. Soc.* 117, 5379 (1995).© 1995, American Chemical Society.

4.2. Ionic Interactions

The recognition of charged ionic species has been a field of intense interest because of their utility in both biology and chemistry [36, 46, 48, 49, 52, 59, 240, 241]. In fact, the recognition of spherical substrates, positively charged metal cations (alkali, alkaline-earth lanthanides) or the negatively charged anions, is one of the simplest recognition processes. From a historical point of view, much of the early work in supramolecular chemistry emerged from efforts to mimic ion recognition in biological systems [240, 241]. In contrast to the large amount of work that has been carried out on cation binding systems and related systems [36, 46, 48, 49, 52, 59], there has been relatively little information on anion binding systems [45, 52]. However, a number of interesting developments have been made in the recognition and sensing of anionic substrates in the last couple of years [45].

4.2.1. Origin and Characteristic Features

Most of the interactions involving these ionic species are predominated by electrostatic interactions. Hence early explanations of the nature of these ionic interactions, which were based on electrostatic models, were very successful in explaining the binding of a large number of charged species to several hosts. However, it was soon realized that in addition to the charge carried by the ionic species, fundamental differences exist between the binding of cations and anions. These differences primarily emerge from the fact that anions have a lower charge to radius ratio than isoelectronic cations (Table 1) [359]. Thus, electrostatic interactions for anions are much less effective than they would be for the smaller cation.

The salient differences between complexes involving cations and anions are excellently illustrated in the small aqueous clusters of alkali metal cations and halide anions (Fig. 17) [360–376]. It should be noted that most of the available experimental investigations of these hydrated ions have benefitted from the pioneering mass spectrometry techniques developed by Okumura and co-workers [377]. A number of theoretical investigations of these hydrated clusters have been carried out to both supplement and understand the experimental investigations [363–369]. They indicate that the cationic clusters are better described than the corresponding anionic clusters. The interaction between water molecules and cations is dominated by ion–dipole interactions at long range. Consequently the hydrogens of the water molecule are oriented away from the ion. At short ranges, electron transfer tends to influence both the geometries and energetics of these systems, which indicates the importance of the size of the cation. Furthermore, the involvement of the oxygen lone pairs in the interaction with

Table 1. A comparison of the radii of isoelectronic cations and anions in octahedral environments.

Cation	r [Å]	Anion	r [Å]
Na^+	1.16	F^-	1.19
K^+	1.52	Cl^-	1.67
Rb^+	1.66	Br^-	1.82
Cs^+	1.81	I^-	2.06

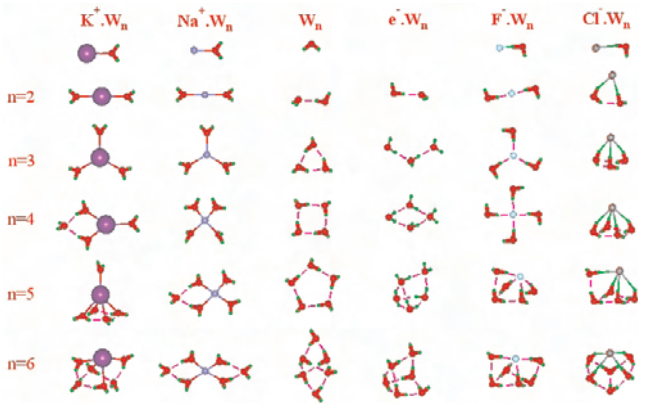


Figure 17. Structures of the complexes of water clusters with anions and cations. For more details, see [225–227, 365–369, 528, 530, 531].

the cation hinders the formation of interwater H-bonding. As a result, most of the cationic water clusters are symmetrically hydrated species. Things are, however, more complicated in the interaction of anions with these water clusters, because the electron clouds of the anion are very diffuse. The larger ionic radii of the halide anions together with the fact that the hydrogens of the water molecule are involved in the binding to the anion leads to an interesting feature of the anionic water clusters [i.e., whether the anions are solvated within a shell of water molecules (internal state) as expected in the condensed phase or the water molecules form a network which then binds the anion in an energetically favorable "surface-solvated" ionic state]. This in a way reflects the competition between the solvation of the anion versus the hydrogen bonding among the water molecules, an issue which will be discussed in more detail subsequently.

In Figure 18, the calculated interaction energies for the binding of a single water molecule with the alkali cations (Na^+, K^+) and the halide anions (F^-, Cl^-) reveal some interesting features. Irrespective of the kind of theory employed, the F^-–H_2O and Na^+–H_2O complexes possess nearly similar interaction energies. But with an increase in the size

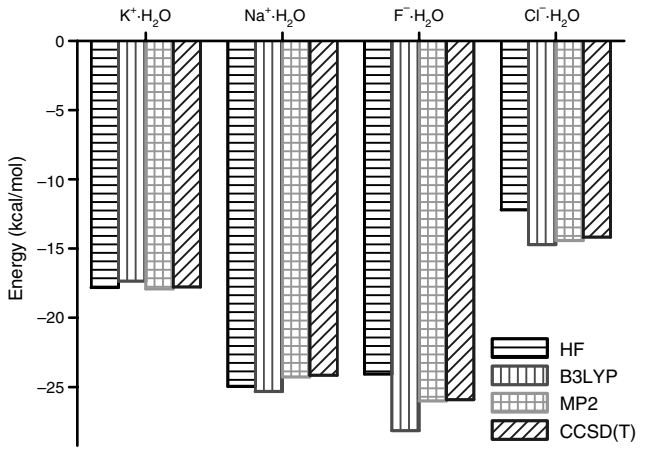


Figure 18. Interaction energies of a single water molecule with different cations and anions. The calculations were carried out using the 6-311++G** basis set.

of the cation and anion, the interaction energy of the K^+–H_2O complex is much larger than that of the Cl^-–H_2O complex. While the reason for the underlying difference is not clear, this observation has important ramifications in the design of novel nanosystems. Interestingly, the nature of the cation (metallic or organic) or anion also influences the binding characteristics. To illustrate this, a plot of the magnitude of all the interaction energy components of five cationic-H_2O complexes is shown in Figure 19. Expectedly, electrostatic interactions are the dominant contributors to the total interaction energy. The magnitudes of the induction and dispersion energies are much smaller with the latter being negligible in the case of the alkali metal cation complexes. Interesting, however, are the magnitudes of the repulsive exchange energies. Even though both the K^+ and NH_4^+ cations have nearly equivalent ionic radii, the magnitude of repulsive energy of the latter is much larger than that of the former. Apart from having an important bearing on the design of novel receptors for these cations, these observations on the intermolecular energetics of these cations explain the large variations in the affinities and selectivities of different receptors [77, 79, 378, 379].

Similar trends observed in the interaction energies of the ion complexes of crown ethers (Fig. 20) dispel the notion that the data obtained on the water monomer complexes would have little relevance to the binding of these cations to more complex systems like macrocycles [221]. It was believed until very recently that the structural rigidity of the macrocycle was one of the most important factors for their high selectivity toward particular cations [56, 86, 196]. However, it soon became evident that the ion–dipole orientation and hence induction energies govern both the magnitude and the equilibrium geometry of the complex [380]. For example, the 12-crown-4 (12c4) ligand, which has S_4 symmetry in the unbound state, has C_4 symmetry in the bound M^+(12c4) state [221]. This indicates that, upon metal complexation, two of the crown oxygens undergo a reorientation so that all four crown oxygens point toward the metal ion, which sits above the plane of the crown by a height that depends on the metal ion radius. Based on the discussion in

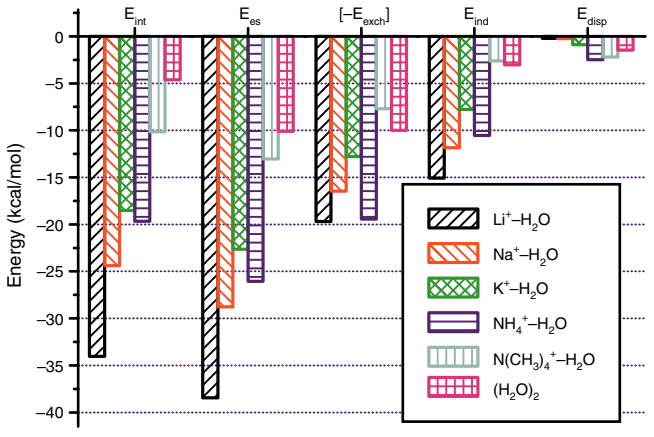


Figure 19. Decomposition of the interaction energies of the cation–water monomer complexes. Note the distinct differences in the binding of metallic and organic cations to water. The calculations were carried out using the 6-311++G** basis set.

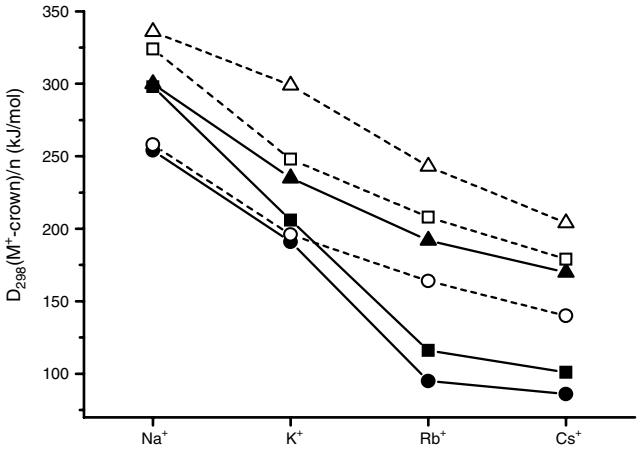


Figure 20. Experimental (closed symbols) and theoretical (open symbols) dissociation energies of M^+(12c4) [circles], M^+(15c5) [squares], and M^+(18c6) [triangles] for M = Na, K, Rb, and Cs at 298 K. Reprinted with permission from [221], M. B. Moore et al., *J. Am. Chem. Soc.* 121, 417 (1999). © 1999, American Chemical Society.

previous sections, this geometrical transformation is a result of the complex trying to maximize the Coulombic and induction energies.

Most of the early receptors, which were designed to bind anions, were positively charged [45, 52, 257–263, 381]. Thus, the binding in most of these systems was mediated through the Coulomb-type electrostatic interactions of the positively charged receptor and the negatively charged anion [45, 381]. However, the efficacy of these positively charged receptors was poor because of problems associated with solvation (i.e., competition for binding between the anionic guests and counterions, etc.) [45, 381]. This led to the development of neutral anion binding receptors [45, 196, 381]. The binding of anions by these neutral receptors is mediated through H-bonding, ion–dipole interactions, or coordination to Lewis acidic centers present in the organometallic ligands. The electrostatic component of the individual interactions in these neutral receptors is relatively smaller than that observed in case of the positively charged receptors. As a consequence, the magnitude of the interaction is strongly dependent on the directionality of the interaction and the electron density of the anion. Apart from these data, there is relatively scarce information on the energetic origin of interaction of large molecular systems containing anions [45]. However, the available data seem to indicate that these interactions are similar to those involving cations albeit with a few exceptions [45].

4.2.2. Manifestation

The early design of cation and anion binding hosts with three-dimensional spheroidal cavities was based on the premise that they would exhibit much stronger binding with the spherically shaped ions than the flat shaped macrocycles [191–193]. Apart from helping magnify several intervening interactions, there is relatively smaller sacrifice in entropy [56, 86, 196]. A more compelling reason for the development of such hosts was the presence of concave-shaped binding sites in biological systems, like the grooves in nucleic acids,

antibody surfaces, neural synapses, etc. [196]. However, it was soon realized that the processes of exchange, regulation, cooperativity, and allostery require an inherent flexibility and adaptivity to change [196]. Thus, the inability of most rigidly organized receptors to respond to these processes far outweighs their benefits of achieving high recognition.

A large number of reviews in the recent past have detailed several aspects of these ionic interactions in various contexts [45, 52, 257–263, 381–390]. Rather than discussing all these aspects, we examine using a few examples how different types of noncovalent interactions involving ions have been useful in the quest for functional nanomaterials. Nearly all the known systems binding ions can be classified as those based on electrostatic interactions, H-bonding, hydrophobicity, and metal ion coordination.

While there is a lot of literature detailing the design of ionophores and receptors based on electrostatic recognition of ions [45, 52, 57, 59, 91, 240, 241], the more interesting aspect of these ionic interactions is their role in stabilizing transition states in enzyme–substrate reactions [254, 391–400]. Thus, there have been several reports of the activation of enzymes in organic solvents by crown ethers [254, 391–400]. The mechanism of this industrially important activation, though unclear, seems to be based on two possibilities [400]. The first is that the crown ether facilitates the transport of tightly bound water molecules from the enzyme active site into the bulk organic solvent and the second is that the crown ether complexes with the surface ammonium groups of lysine [254]. The available experimental evidence seems to validate the latter mechanism [254]. The macrocyclic interactions lead to a reduced formation of inter- and intramolecular salt bridges and consequently lower the kinetic conformational barriers. This enables the enzymes to refold into thermodynamically stable, catalytically (more) active conformations. Ionic interactions also play a direct role in the catalytic enhancement observed in enzymes. In this context, one encounters the concept of a short strong H-bond (**D**-H$\cdots$**A**$^-$), an H-bond which is characterized by high interaction energies (>15 kcal/mol) and very short H-bond distances [401–413]. In Figure 21, we highlight the role of these short strong H-bonds in explaining the origin of the spectacular catalytic enhancement observed in case of the enzyme Ketosteroid isomerase [146, 412, 413].

The work of Ward and co-workers, wherein the salts formed from organodisulfonate and guanidnium ions have been used to design an extremely large number of complex molecular architectures, is the one of the most elegant examples of the use of H-bonding based on ionic interactions in the custom design of nanomaterials [55, 414–417]. Using a variety of substitutents and strategies, a bewildering array of crystal architectures was designed (Fig. 22). Interestingly, these molecular frameworks formed as a result of these weak noncovalent interactions are structurally similar to the zeolites, wherein the metal-oxygen bonding yields a more robust and strong framework [418, 419]. The work of Seddon and Aäkeroy on the development of novel crystal architectures using various organic cations and anions also merits mention in this context [420].

The synthesis of helicates (Fig. 23) is a nice example of metal ion mediated recognition because the ligands (analogs of quatepyridine) are spontaneously assembled in the presence of Cu^+ ion [421, 422]. The double-stranded helicates are formed as a result of the ligand strands wrapping around the metal ions in a double helical pattern [421, 422]. Of particular importance is the fact that the structures are a manifestation of the tetrahedral-like coordination imposed by each Cu(biypyridine)$_2^+$ unit [421, 422]. Other metal ions like Ag^+ have also been employed in synthesis of some of the variants of these helicates [423]. The self-assembly of these helicates brings forth the issue of positive cooperativity, wherein the complexation with the metal makes the second complexation easier, and so on. Furthermore, self-recognition is also an important characteristic for the generation of these architectures because in a mixture of ligands, a helicate is preferentially consituted by identical ligands [424]. While these double-stranded dihelicates were based on metal cations, the circular helicates result from assembling Fe^+ coordinated ligands around a chloride ion [424].

Most of the ionic interactions involving the use of hydrophobic effects to trigger the recognition phenomena

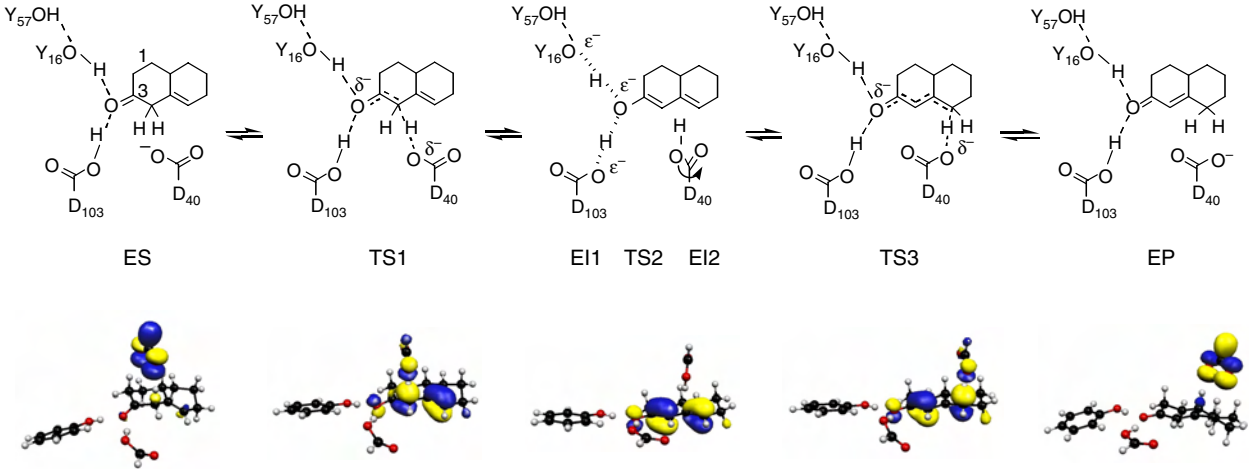


Figure 21. Role of ionic interactions in stabilizing enzyme-substrate transition states. Enzymatic reaction scheme of ketosteroid isomerase. The negatively charged amino residues in the active site of the enzyme are crucial to its mechanism. Reprinted with permission from [146], K. S. Kim et al., *Proc. Natl. Acad. Sci. USA* 97, 6373 (2000). © 2000, National Academy of Science.

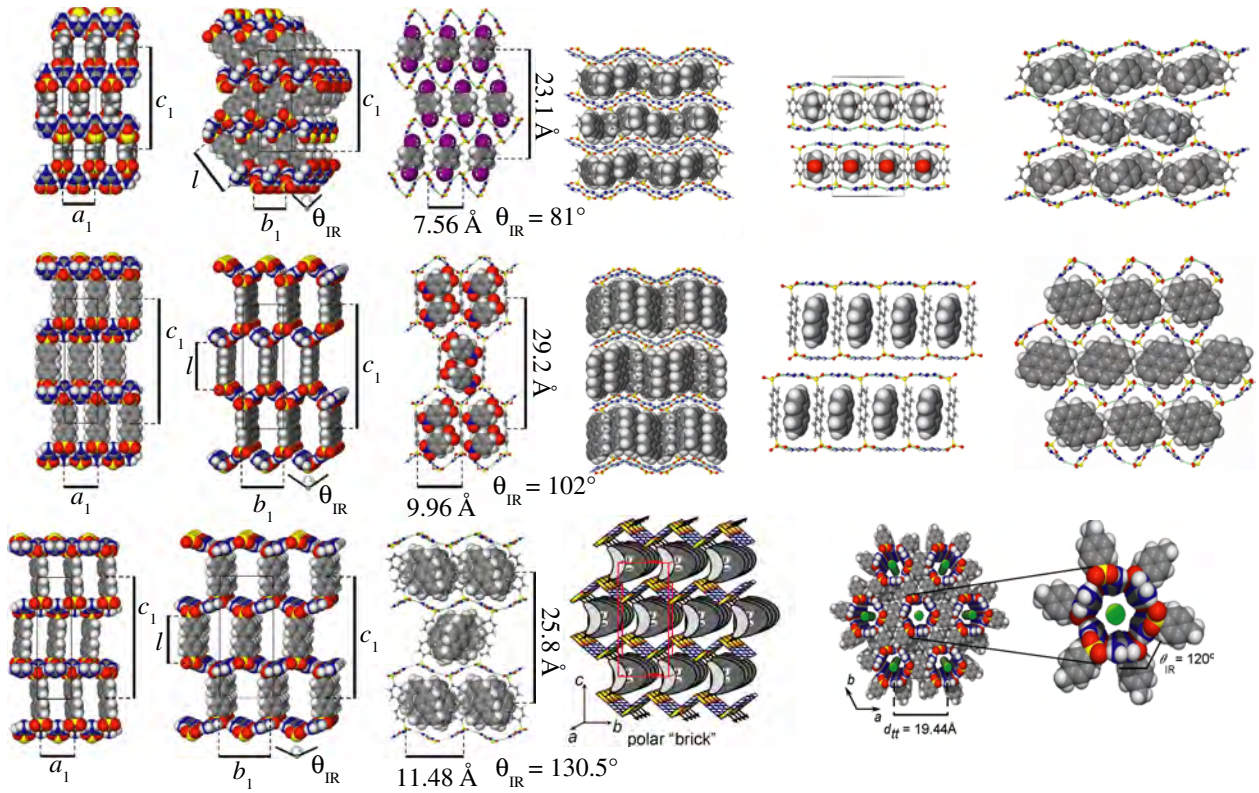


Figure 22. Novel molecular architectures formed from the salts of organodisulfonate and guanidinium ions. Note that some of these architectures also host guest molecules. For more details, see [55, 414–417].

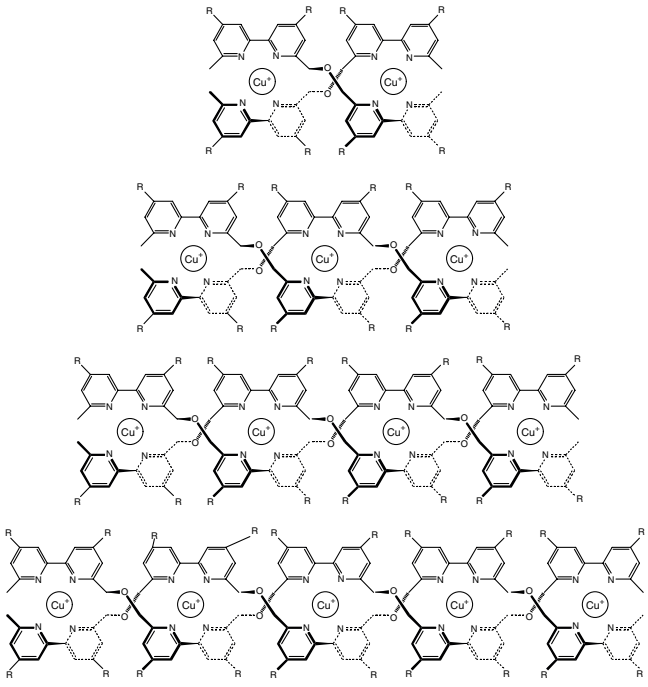


Figure 23. Schematic representation of metallohelicates obtained from the reactions of oligobipyridines with Cu^+ ion. Reprinted with permission from [59], V. G. Machado et al., *J. Braz. Chem. Soc.* 12, 431 (2001). © 2001, Brazilian Chemical Society.

involve π systems [425], which is the focus of discussion in the next section. For example, the inclusion of naphthalenesulfonates by β-cyclodextrin is a result of the displacement of the water molecules from the internal cavity of cyclodextrin by the naphthalene residues. The anionic sulfonate is in contact with the solvent and thereby controls the orientation of the naphthalene residues in the cavity [425].

4.3. Interactions Involving π Systems

Among the large number of noncovalent intermolecular interactions discussed in this chapter, interactions involving π systems are one of the principal noncovalent forces governing molecular recognition [4, 7, 43, 53, 54, 56, 61, 69, 87, 88, 353, 426–437]. They influence the structures of proteins, DNA, host–guest complexes, solid materials containing π systems, crystal structures and self-assembled supramolecular architectures [4, 43, 53, 54, 56, 61, 69, 87, 88, 353, 426–437]. These interactions also control the intercalation of certain drugs into DNA [51, 167]. The conductive properties of molecular wires formed by self-assembly of stacks of aromatic macrocycles are also of recent interest [10, 121]. A distinct advantage of the use of the interactions involving π systems in the fabrication of nanodevices is the fact that the subtle changes in the electronic characteristics of the π system can lead to dramatic effects on the structure and properties of the nanosystem. Furthermore, these changes in the electronic characteristics can be controlled either by the electrochemical or photochemical means.

However, an understanding of the nature of their origin has only started being unravelled recently. One of the reasons for this is that the strength of most of these interactions is relatively small and is dominated by dispersion interactions. As was mentioned in previous sections, it is difficult both theoretically and experimentally to quantify the contribution of dispersion interactions to the total interaction energy.

A few studies which have quantified the strength of these interactions involving π systems indicate that the strength is dependent both on the nature of the π system and the kind of countermolecule interacting with it [232, 438–453]. In general, the strength of interactions involving cations and π systems is greater than the other interactions involving π systems (π–π, π–H, π–Lewis acid). Given the ubiquity of these interactions involving π systems in a number of areas in nanoscience and nanotechnology, a detailed description of these interactions involving π systems is presented in the following sections.

4.3.1. Classification

One can classify the interactions involving these π systems based on the nature of the countermolecule with which they are interacting. Thus, when the countermolecule is a metal cation, electrostatic and induction energies dominate the interaction energy. Dispersion interactions dominate when the countermolecule is either a rare gas atom or a nonpolar molecule (gas dimers, hydrocarbons). A combination of electrostatic and dispersion interactions governs the interaction energy, when the countermolecule is polar (water or methanol clusters, Lewis acids). Unlike other kinds of interaction discussed earlier, the magnitude of repulsive interactions plays a critical role in governing the observed geometry.

As would be shown in subsequent sections, the nature of the π system (olefinic, aromatic, heteroaromatic) also influences the characteristics of these interactions. Even in aromatic systems, the presence of electron donating substituents leads to complexes with enhanced interaction energies, when compared to benzene. This enhancement is due to an increase in both dispersion and electrostatic interactions. On the other hand, the geometries of complexes formed with aromatic rings having electron withdrawing substituents is extremely dependent on the characteristics of the interaction energy. Thus, when the interaction energy is dominated by dispersive interactions, a π type of complex is formed. When electrostatic energies predominate the interaction energy, a σ type of complex involving the electron withdrawing substituent is formed. When the countermolecule is a water or methanol cluster, the size of the cluster also influences the nature of the interaction with the π system, resulting in a crucial influence on the nature of the resulting structure (σ or π).

Since most of the interactions involving π systems are mediated through the negatively charged π electron cloud, the early studies were of the opinion that purely electrostatic considerations could explain the observed trends of the interaction energies. However, it was soon realized that a purely electrostatic model could not explain the binding of rare gases or neutral molecules like methane or benzene to these π systems. Furthermore it could also not explain the exceptionally high binding energies observed in the case of the ethene systems. For example, the most recent experimental estimates of the enthalpies of binding (ΔH_0) of Na^+ to ethene and benzene are 10.7 ± 1.0 and -21.5 ± 1.0 kcal/mol, respectively [454]. Even though geometrical considerations are taken into account, the general impression would be that the enthalpy of binding of Na^+ to benzene to be three times that of ethene since benzene has three times the number of π electrons. A more interesting problem arises when one compares the interaction energy of the benzene–Na^+ and the pyrrole–Na^+ complexes. Though both of them have an equal number of π electrons, the cation–π interaction in the latter is stronger than in the former. It therefore becomes apparent that one has to take into account the differential electron density distributions of these π systems to explain the trends in interactions. In terms of energetic jargon, it implies that the magnitudes of the individual interaction energy components have to be taken into account. In the following account, we show some of the characteristics of the interactions of these π systems with various countermolecules.

4.3.2. Origin and Characteristic Features

We begin our discussion of the energetic origin of these π complexes with the cation–π interaction [432, 450]. The early explanations of origin of the cation–π interaction were predominantly based on electrostatic considerations [432]. Thus a strong cation–π interaction was associated with a more negative electrostatic potential at the center of the π system [432]. However, after a study of the interaction of the Na^+ cation with different aromatic π systems, it was realized that the variations in the electrostatic potential of the π system only mirrored the trends in the cation–π interaction energies [432]. One has to take into account other factors like the polarizability of the π system and the nature of the cation to obtain accurate estimates of the interaction energies [432]. In this connection, it is interesting to note that a detailed theoretical investigation of the interaction of the tetramethyl ammonium cation with benzene highlighted the importance of dispersion energies in the cation–π interaction [438, 439].

In order to understand these cation–π interactions, it is useful to examine the results obtained from a recent and extensive study of the interaction of different cations, which include both metallic (Li^+, Na^+, K^+, Ag^+) and organic {ammonium (NH_4^+), guanidinium ($C(NH_2)_3^+$), tetra methyl ammonium ($N(CH_3)_4^+$)} cations, with different classes of π systems, viz. alkenes (ethene), arenes (benzene), and heteroarenes (pyrrole) [450]. It can be seen from Figure 24 that the electrostatic and induction energies are the major contributors to cation–π interactions involving alkali metal cations. In particular it can be noted that the magnitude of the induction energies is nearly double that of the electrostatic energies. This predominance of the induction energies can be attributed to the interaction between the highest occupied molecular orbital (HOMO) of the π system and the empty *s* orbital of the metal cation. This MO interaction is in turn governed by energies of the orbitals involved and the size of the cation, because a smaller sized cation can

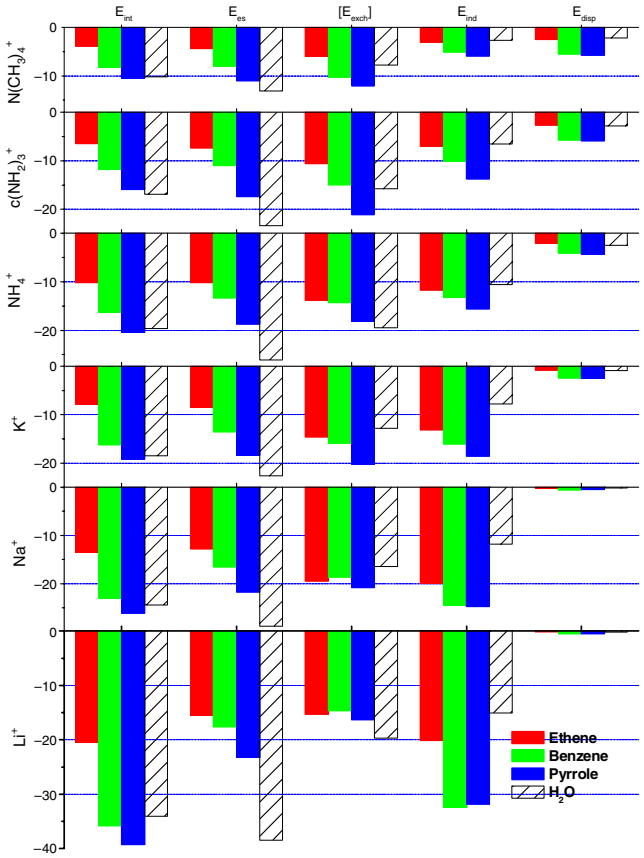


Figure 24. Decomposition of the interaction energies of cation–π and cation–water complexes. Note the distinct differences between the interaction of metallic and organic cations with various π systems. The calculations were carried out using the 6-31 + G* basis set. Reprinted with permission from [450], D. Kim et al., *J. Phys. Chem. A* 107, 1228 (2003). © 2003, American Chemical Society.

have a closer approach to the π system due to a decrease of exchange–repulsion at a given separation. The smaller intermolecular separation leads to a larger electrostatic stabilization and also facilitates an enhanced overlap of the constituting MOs, which in turn leads to enhanced induction energies. It should be noted that several authors attribute this kind of MO interaction to be of the charge-transfer type. The contribution of the dispersion energies is quite significant in case of the K^+ complexes, which is in contrast to the neglibile magnitudes observed in the Na^+ and Li^+ complexes [450].

The pyrrole complexes exhibit the largest electrostatic energies because the π electron density of pyrrole is much higher than that of either benzene or ethene [450]. This is because the nitrogen lone pair is a part of the delocalized π system of pyrrole [455]. The presence of the nitrogen atom in pyrrole, however, influences both the equilibrium geometries and the interaction energies exhibited by these pyrrole complexes. While the magnitudes of the induction energies are nearly similar in the Na^+ and Li^+ complexes of both benzene and pyrrole, it is much higher in case of the pyrrole–K^+ complex as compared with the benzene–K^+ complex. The innate tendency of both the small-sized Li^+ and Na^+ cations is to maximize the electrostatic stabilization by a closer approach to the π system. However, the presence of the electronegative nitrogen atom prevents a closer approach of the cation to the π system, as can be noted from the magnitude of the exchange energies (Fig. 24). The resulting larger intermolecular separation in the pyrrole complexes, as compared to the corresponding benzene complexes, limits the extent of orbital overlap. Consequently the magnitudes of the induction energies are nearly similar in the Li^+ and Na^+ complexes of benzene and pyrrole. On the other hand, the larger size of the K^+ cation precludes a close approach to either benzene or pyrrole because of increased exchange–repulsion. Consequently, both the benzene and pyrrole complexes exhibit nearly similar intermolecular separations. However, the enhanced π electron density of pyrrole results in increases in the magnitudes of both the induction and electrostatic energies.

The magnitudes of the electrostatic and induction energies in the case of the organic cation complexes of these π systems are much smaller than those observed in case of the π–alkali metal cation complexes (Fig. 24). While the repulsive exchange energies are relatively smaller, the dispersion energies exhibit large increases. Thus, a comparison of the pyrrole complexes of Li^+ and $N(CH_3)_4^+$ cations reveals that the dispersion energies of the latter are nearly 12 times larger than that of the former. However, the total interaction energy of the pyrrole–Li^+ complex is only 4 times larger than that of the pyrrole–$N(CH_3)_4^+$ complex. Electron correlation plays a very important role in the latter interaction because the π–σ^* interaction (where π is the HOMO of the π system and σ^* is the LUMO of the organic cation), which is of the inductive type, is magnified [438, 439, 450]. One of the reasons for the magnification of this induction energies upon inclusion of electron correlation is the fact that there is a depletion of electron density from the centers of the bonds and a concurrent increase in the shells around the atomic nuclei [456]. As a result, a much closer approach of the cation to the π system is facilitated. This results in an increased MO interaction and as a result a larger induction energy. The increase in the electron density around the shells of the atomic nuclei and the closer approach of the cation to the π system result in the increased dispersion energy. As in case of the interaction of these π systems with the alkali metal cations, the electrostatic energies are largest in the case of the pyrrole containing complexes. However, unlike what is observed in the case of the alkali metal cations, the induction energies are much larger in the case of the pyrrole containing complexes. Given the nearly similar intermolecular separations and the equal number of delocalized π electrons in both benzene and pyrrole, the dispersion energies are nearly similar in their organic cation complexes.

Given our discussion of the cation–water interactions in an earlier section, it is of interest to compare their characteristics with the cation–π interactions. It can be seen from Figure 24 that the total interaction energies of the cation–water complexes are comparable in magnitude to the corresponding cation–pyrrole complexes. However, what distinguishes the interactions of these cations with water and the π systems is the magnitude of the electrostatic energies, with them being dominant contributors to the total interaction energy in case of the former. The marked difference

between the water and π complexes of these cations can be attributed to the difference in the nature of the donor. In water, the donor lone pair of the oxygen atom is more localized than the diffuse π cloud of the π system. As a result, the polarizability of the former is much smaller than the latter. Hence the electrostatic contributions are larger and the induction energy contributions are smaller in case of the water complexes. The repulsive energies are higher in the case of the water complexes because the cation encounters another atom in contrast to a bond in ethene or the center of the aromatic ring (no atoms or bonds). In the context of the design of receptors, it is useful to examine the implications of these findings in understanding chemical and biological processes. By definition, a greater electrostatic contribution to the interaction energy of the complex implies that the magnitudes of the interaction energies are susceptible to the dielectric of the solvent. The fact that the electrostatic contributions are much lower in the case of the cation–π complexes than in the corresponding cation–H_2O complexes might explain the wide prevalence of the former in chemical and biological systems.

The interaction of these π systems with transition metal cations like Ag^+ requires special attention owing to the large interaction energies and the presence of *d* electrons (Table 2). The small size of the Ag^+ ion (1.26 Å) implies that the Ag^+ cation can approach much closer to the π system. However, the larger nuclear charge of the Ag leads to a much larger repulsive energy. The higher dispersion energy results from both a smaller intermolecular separation and the participation of the *d* electrons in the binding. The higher interaction energies also imply that the role of the π system in the binding is minimal. However, the equilibrium geometries can be expected to vary owing to the widely differing magnitudes of the various interaction energy components.

Since most instances of cation–π interactions in chemical and biological systems exhibit significant deviations from the kind of idealized behavior discussed, it is of interest to examine on how the magnitudes of the various interaction energy components are modulated when the cation is brought closer to the π system from larger distances. In order to do so, we have compared the interaction energy components of the benzene–NH_4^+, benzene–K^+, and ethene–K^+ complexes for various intermolecular separations in Figure 25. While the former two complexes were chosen because of the differences in cations, the latter was chosen because it allows us to examine the role of the π system. At large intermolecular separations, electrostatic energies dominate the interaction

Table 2. Interaction energy components of Ag^+–π complexes.

	Ethene	Benzene	Pyrrole
E_{int}	−28.2	−29.9	−40.3
E_{corr}	−7.3	−10.5	−10.5
E_{elec}	−44.8	−40.9	−52.5
E_{exch}	139.2	96.7	124.2
E_{ind}	−110.4	−70.6	−98.6
E_{disp}	−12.1	−15.2	−13.4

Note: All energies are in kcal/mol. For more details of the theoretical methodology employed, see [450].

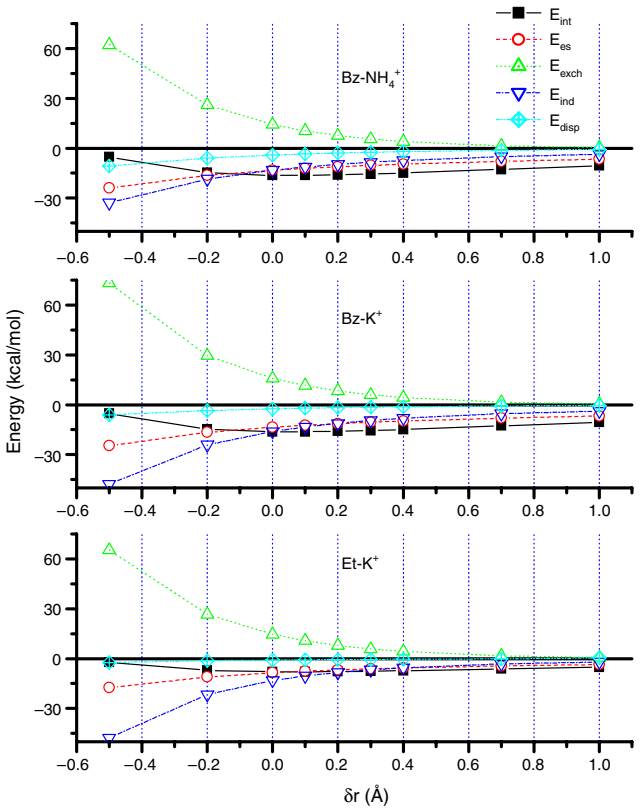


Figure 25. Variation in the magnitudes of intermolecular energy components of benzene–NH_4^+, benzene–K^+, and ethene–K^+ complexes with changes in the intermolecular distance. The calculations were carried out using the 6-31+G* basis set. Reprinted with permission from [450], D. Kim et al., *J. Phys. Chem. A* 107, 1228 (2003). © 2003, American Chemical Society.

energies of these complexes. Consequently the interaction energies for the benzene–K^+ complex are much larger than those observed in the ethene–K^+ complex but similar to those observed in the benzene–NH_4^+ complex. At short intermolecular separations, induction energies become dominant in the metal cation complexes with both the ethene–K^+ and benzene–K^+ complexes exhibiting nearly similar induction energies. The predominance of the induction energies over that of the electrostatic energies can be observed long before the onset of the equilibrium geometry in the metal cation complexes. Unlike what is observed in the metal cation complexes, much larger increases in the dispersion energies are observed in the benzene–NH_4^+ complex when the intermolecular separation is decreased. The repulsive exchange energies are much smaller in the benzene–NH_4^+ complex than in the corresponding benzene–K^+ complex and interestingly are similar in magnitude to those observed in the ethene–K^+ complex. Thus, the nearly similar interaction energies of the benzene–NH_4^+ and benzene–K^+ complexes result from a balance of dispersion and induction energies because the electrostatic and exchange energies are nearly similar and hence cancel out. The predominance of the dispersion energies in the case of the benzene–ammonium complexes implies that suitable receptors specific for the NH_4^+ cation could be designed by maximizing the dispersion energies [378, 379].

An interesting variant of the cation–π interactions is the recently discovered anion–π interactions [457–461]. Most of the studies have devoted their attention to the interaction of an electron deficient π system like hexafluorobenzene with an anion [457–459]. Though there have been no studies on the origin of the interaction, the available evidence seems to suggest that inductive energies dominate the interaction. Though there are reports of electron rich π systems like benzene interacting with oxygen or water cluster anions, the π system does not seem to play any role, with the interaction being primarily mediated through the benzene hydrogens [462].

In the order of decreasing magnitude, the interaction energies of the π–Lewis acid complexes follow the cation–π interactions. Apart from a few experimental reports detailing the interaction of Lewis acids with π systems in the context of donor–acceptor complexes and electrophilic aromatic substitution reactions, little was known about either the magnitudes or the exact nature of their interaction [463, 464]. However, in the course of our extensive investigations, it was found that the halides of aluminum, boron, and hydrogen form fairly strong noncovalent complexes with both benzene and ethene [441, 465, 466]. The magnitude of the interaction depends on the nature of the Lewis acid, with the benzene complexes of aluminum halides exhibiting interaction energies (8–15 kcal/mol) which are significantly larger than the corresponding complexes of the boron halides (3–6 kcal/mol) [441, 465, 466]. As far as our knowledge goes, the interaction energies exhibited by the aluminum halide complexes of benzene are the highest for a neutral (uncharged species) binding to it. A major difference in the binding of benzene or ethene with boron and aluminum halides is that in the complexes of the latter, electrostatic interactions have a dominant contribution to the interaction energy, while in the complexes of the former, dispersive interactions have a dominant role (Fig. 26). These findings seem to indicate that the aluminum halide complexes conform to the general behavior of donor–acceptor complexes [463]. One of the most interesting consequences of the interaction of Lewis acids with benzene is that in the lowest energy structure, the carbon atom of benzene which lies directly below the Lewis acid exhibits a dramatic increase in the negative charge when compared to the remaining five carbon atoms. A simultaneous increase in the positive charge on the hydrogen atom attached to this carbon is also noted. This increase in the negative charge (or nucleophilicity) of the carbon atom, which is independent of the nature of the Lewis acid interacting with benzene, could ease an electrophilic attack on this particular benzene carbon as a result of a significant lowering in the activation energy. Furthermore the increase in the positive charge on the hydrogen attached to this carbon would facilitate its departure after an attack by an electrophile. Similar effects have been noted in the interaction of some of the transition metal cations with benzene and C_{60} [435, 463]. The implication of this observation is enormous, because to date, the role of Lewis acids in electrophilic aromatic substitution reactions was perceived to be limited to the generation of the electrophile [464].

Complexes of the first-row hydrides (CH_4, NH_3, H_2O, HF) with π systems (π–H interaction) are of particular interest because they exhibit interactions [447] which are the prototypes of a typical hydrogen bond. In Figure 27, we

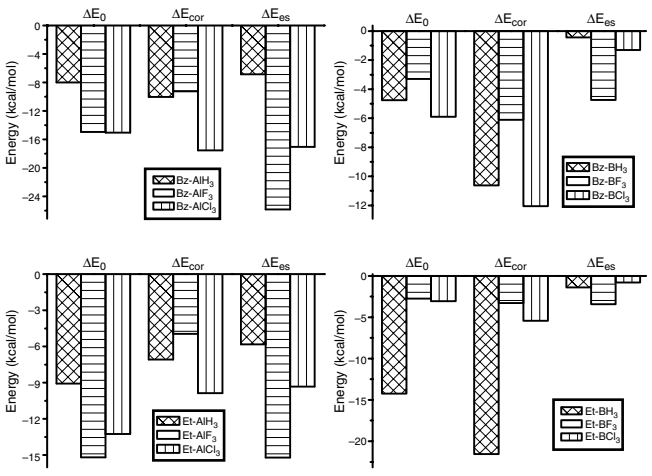


Figure 26. Contribution of electrostatic and dispersion energies in the π–Lewis acid complexes. The calculations were carried out at the MP2/6-311++G** level of theory. Reprinted with permission from [232], K. S. Kim et al., *Chem. Rev.* 100, 4145 (2000). © 2000, American Chemical Society.

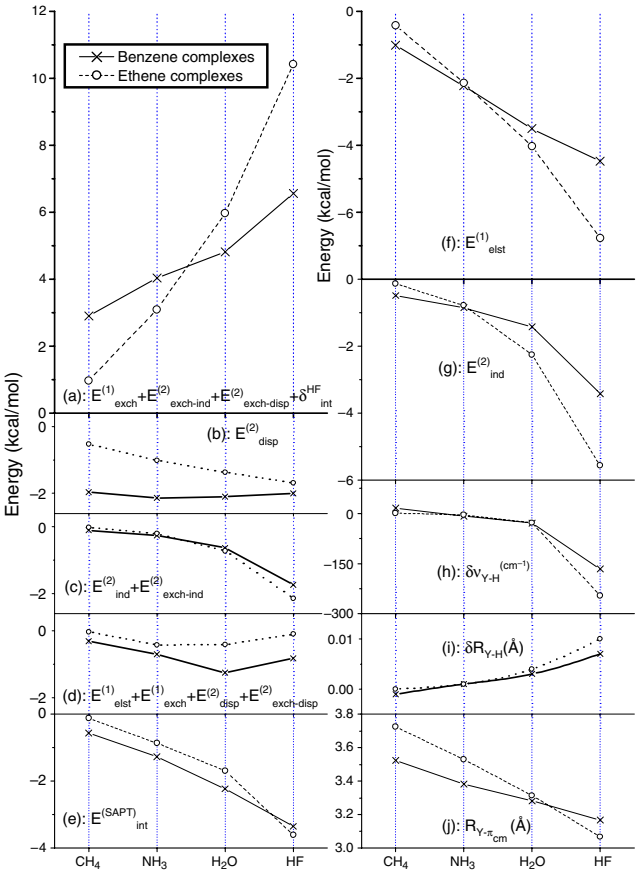


Figure 27. Trends in the interaction energy components, geometrical parameters, and spectral features of the π–H interaction. The calculations were carried out using the 6-31+G* basis set. Reprinted with permission from [447], P. Tarakeshwar et al., *J. Am. Chem. Soc.* 123, 3323 (2001). © 2001, American Chemical Society.

highlight the trends of various interaction energy components, their sums, spectral shifts, and selected geometrical parameters of various first-row hydride complexes of ethene and benzene complexes. It can be noted from Figure 27a that the increase in the repulsive exchange energies is more pronounced in the ethene complexes than in the benzene complexes. The reason for the anomalous behavior of the exchange energies in case of the ethene and benzene complexes lies in both the size of the donor (hydride) and the way it approaches the π system. As we progress from CH_4 to HF, the decrease in the electron density of the H-atoms involved in the π–H interaction leads to a closer approach to the π system because of a decrease of exchange repulsion at a given separation. The widely differing electron density profiles of benzene and ethene implies that the hydride aims at the midpoint of the ring (where there are no nuclei or bonds) in benzene. Consequently the decrease of the intermolecular $R_{Y-\pi_{CM}}$ distance with the decrease of the donor size would be smaller for the benzene than for the ethene complexes (Fig. 27j). The smaller variation of $R_{Y-\pi_{CM}}$ in the benzene complexes also explains the near constancy of the dispersion energies (Fig. 27b). On the other hand, the somewhat smaller decrease of $R_{Y-\pi_{CM}}$ in the ethene complexes, as we progress from NH_3 to HF, explains the slight increase in the dispersion energies (Fig. 27b). A comparison of the electrostatic and induction components of the ethene and benzene complexes plotted in Figure 27f and g reveals that in the case of complexes involving CH_4 and NH_3, their magnitudes are comparatively smaller for the ethene complexes. However, with the progressive decrease in the hydride size, the magnitude of these components in the ethene complexes is larger than the magnitudes observed in the corresponding benzene complexes. It is instructive to note that this behavior of the attractive induction energies was also observed in case of the cation–π interactions [450].

Since the donor is a diffuse π electron cloud in the case of a π H-bond, and the lone pair of an electronegative atom in the case of a conventional H-bond, it is useful to examine their salient similarities and differences. In Figure 28, the magnitudes of the interaction energy components of the water dimer and the ethene–water dimer at different intermolecular separations are plotted. It can be seen that in the lowest energy structure of these complexes, the replacement of a single water molecule by ethene results in a decrease of the electrostatic contribution of the total attractive energy by about 20% [447]. Expectedly, the contributions of induction and dispersion energies to the total attractive interaction of the water dimer are much lower than in case of the ethene–water dimer. In both systems, the magnitude of the repulsive exchange energies has a far more important role in the characterization of the equilibrium geometry than the magnitude of the attractive energies like electrostatic, inductive, and dispersive energies.

The role of dispersion and repulsive energies in governing the geometries of these π complexes is well illustrated in a study of the argon complexes of fluorine-substituted benzenes [448]. These complexes are interesting because irrespective of the widely different electron densities of the π systems, they all exhibit similar binding energies [448]. Furthermore, in all the cases, the rare gas is located on top of the π system. A detailed theoretical investigation revealed

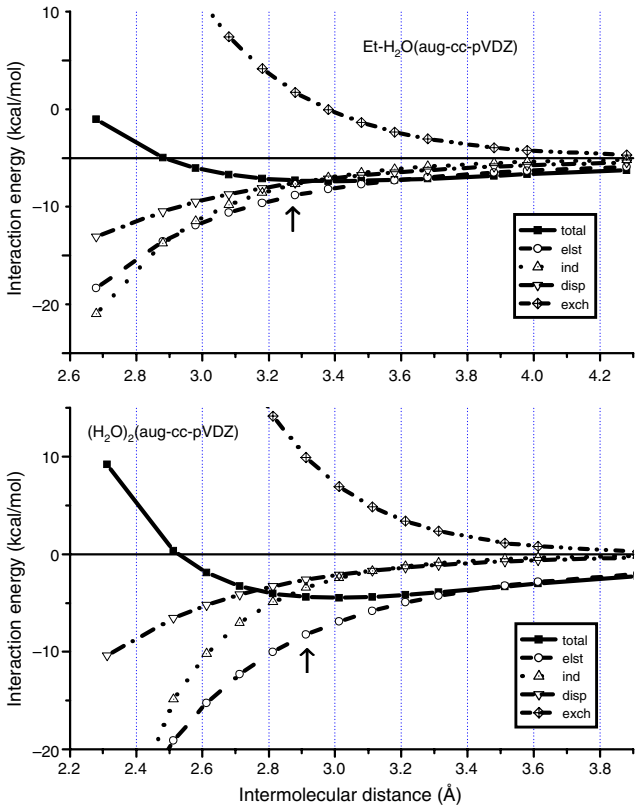


Figure 28. Variation in the magnitude of various interaction energy components as a result of changes in the intermolecular distance of the ethene–H_2O and $(H_2O)_2$ complexes.

that the lower π electron density of the *p*-difluorobenzene facilitates a much closer approach of the argon atom to the center of the π system because of lower exchange–repulsion. The dispersion energies are, however, maximized because of the decrease in intermolecular separation [448]. The role of dispersion energies can also be noted in the benzene complexes of paramagnetic atoms like nitrogen [449]. We discuss the characteristics of these paramagnetic complexes in a subsequent section.

The π–π interactions are one of the most intriguing noncovalent interactions, in the sense that the negatively charged and diffuse electron clouds of the π systems exhibit an attractive interaction [43, 87, 437, 451–453, 467]. The interaction is predominated by dispersion interactions, when the π systems possess similar electron density. However, when one of the systems is electron rich (benzene) and the other electron deficient (hexafluorobenzene), the resulting complexes are bound by induction interactions with the negative charge being transferred from benzene to hexafluorobenzene [468, 469]. The acetylene dimer is one of the simplest systems exhibiting a π–π interaction [470–473]. Though a parallel displaced conformer which involves a stacking of the two acetylene π systems exhibits a π–π interaction, the most stable conformer is a T-shaped structure, which involves the interaction of the acetylinic hydrogen with the π system of the other [473]. In the case of the ethene dimer, the most stable structure is stacked D_{2d} structure (Fig. 29), which involves the formation of a

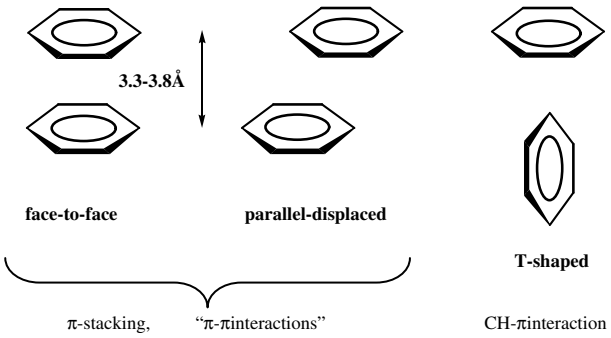


Figure 29. Different orientations of π–π interactions.

quadruple hydrogen bond [474, 475]. The interaction of two benzene rings (benzene dimer) has been widely investigated both experimentally and theoretically [437, 451–453, 467]. The experimental estimates of the interaction energy seem to be around ~2 kcal/mol, which indicates that the attraction is appreciable and significantly influences the interaction of phenyl rings in solution or other environments, in addition to other factors such as solvophobic effects [56, 196]. The interaction is predominantly based on dispersion interactions and as in case of the acetylene dimer, the benzene dimer can manifest itself in any one of the structural forms: stacked, T-shaped, or parallel-displaced (Fig. 29) [437, 467]. The available evidence seems to indicate that the parallel displaced conformation is most stable [451–453]. However, it should be noted that the isolated benzene dimer is highly fluxional and can coexist in any of the three structural forms.

A principal advantage of π–π interactions in the context of nanorecognition is the fact that the interconversion between the different conformers or the orientations of the two interacting π systems (Fig. 30) can be exquisitely controlled [230, 437, 467]. While the use of this control in the design of functional nanodevices is described in subsequent sections, it is of interest to pictorially examine the nature

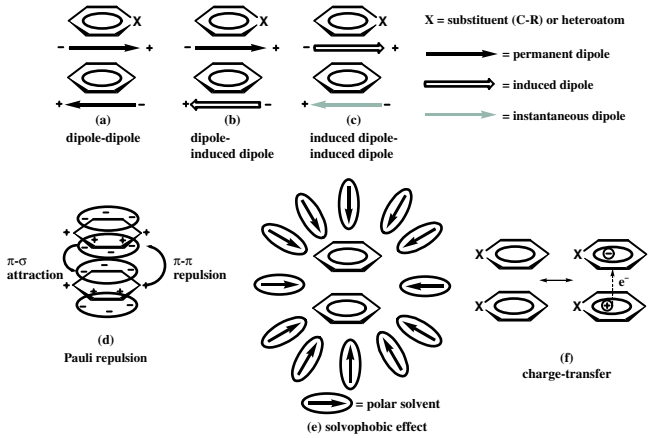


Figure 30. Schematic illustration of the electrostatic interactions between two arene moieties. For simplicity the two arene moieties are drawn in perfect face-to-face alignment. Reprinted with permission from [467], C. Janiak, *J. Chem. Soc., Dalton Trans.* 21, 3885 (2000). © 2000, Royal Society of Chemistry.

of these π–π interactions in terms of the individual components (Fig. 30) [467].

(a) Dipole–dipole (electrostatic) interactions: the interactions between the different permanent and static molecular charge distributions.
(b) Dipole–induced-dipole interactions between the static molecular charge distribution of group **A** with a proximity-induced change in charge distribution of group **B**.
(c) Induced-dipole–induced–dipole (London) dispersion interaction: the instantaneous dipole moment from a fluctuating electron cloud polarizes a neighboring molecule and induces the charges, resulting in an instantaneous dipole.

These van der Waals interactions (a)–(c) are inherently attractive and their potentials fall off rapidly with distance by $1/r^6$.

(d) Pauli repulsion: at very short distances the filled electron clouds of the interacting molecules begin to overlap and the (Pauli) repulsion between the electron shells becomes dominating.
(e) Solvophobic effects: in polar solvents, solvophobic effects leading to desolvation can stabilize an aggregation of lipophilic molecular surfaces.
(f) Charge transfer: charge transfer or electron donor–acceptor effects refer to a stabilization due to mixing of a ground state $\mathbf{A}\cdots\mathbf{B}$ with an excited charge-separated state $\mathbf{A}^+\cdots\mathbf{B}^-$. Such Coulomb attractions based on charge transfer are generally viewed as negligible and are different from static-electron donor–acceptor complexes.

The only remaining interactions which need to be discussed in the context of interactions involving π systems are the interactions involving two π systems or the interaction of a π system with neutral molecules like dihalogens [232, 463]. The complexes of these π systems with dioxygen and dihalogens have been of interest in the context of understanding gas–surface reactions and as models of charge-transfer systems [463]. In most cases, the interaction in these π complexes is dominated by dispersion energies [453, 476]. However, in case of the benzene–I_2 complex, induction energies also play a role in the binding [477].

4.3.3. Manifestation

The literature on the use of intermolecular interactions involving π systems in the design of ionophores, nanomaterials, and nanosystems is almost impossible to summarize, but an effort will be made to demonstrate their omnipresence on the basis of a few examples [7, 43, 53, 54, 56, 61, 69, 87, 88, 353].

A large number of ionophores like cyclophanes, calixarenes, etc. are based on the interaction of the ions with the π systems present in these ionophores [196, 250]. The cation–π interactions have been invoked to explain the mechanism of ion channels and more recently in understanding substrate and/or transition state recognition of several enzymes like phospholipase C, thrombin, DNA glycosylases, etc. [53]. Recently, there also has been a report of the use of cation–π interactions in stabilizing α-helices of proteins in water

[187]. An area which has benefited the most from detailed investigations of cation–π interactions is the design and development of chemsensors [50, 54]. Thus, a large number of systems capable of selectively binding microscopic traces of toxic or biologically important ions have been developed [50, 54]. Figure 31 shows the tripodal class of receptors, where suitable changes in the nature of the ligand (L) leads to highly specific binding of biologically important ions like ammonium cation, acetylcholine, etc. [77, 79, 378, 379]. Though the mechanism of ion binding is different, this tripodal class of receptors has also been found to be useful in binding biologically important anions like chloride, fluoride, etc. [383]. In the context of nanomaterials, the recent use of cation–π interactions in the synthesis of silver grid-type metalloarchitectures, C_{60}–Ag^+ interactions, and ultrathin silver nanowires also merits mention [384, 435].

The earliest expression of π–H interactions was in the "molecule-within-molecule" complexes of Cram and Collet, carcerands and cryptophanes [478]. An interesting and recent example of the use of the π–H interactions is in the encapsulation of hydrocarbon guests within the cavernous environs of the resorcinarene-based nanoscale containers (Fig. 32) [479]. An area wherein π–H interactions are expected to have a tremendous impact is the interface between biological molecules and novel nanomaterials like carbon nanotubes, fullerenes, etc. [478–480]. An understanding of this interface is central to the design and development of sensors and miniature devices for a wide variety of biological applications [481]. The interaction of water with carbon- and silicon-based surfaces is also attracting a lot of interest because of its implications in the development of carbon- and silicon-based device technologies.

π–π interactions are widely prevalent in nature. However, their use as a building tool in supramolecular chemistry

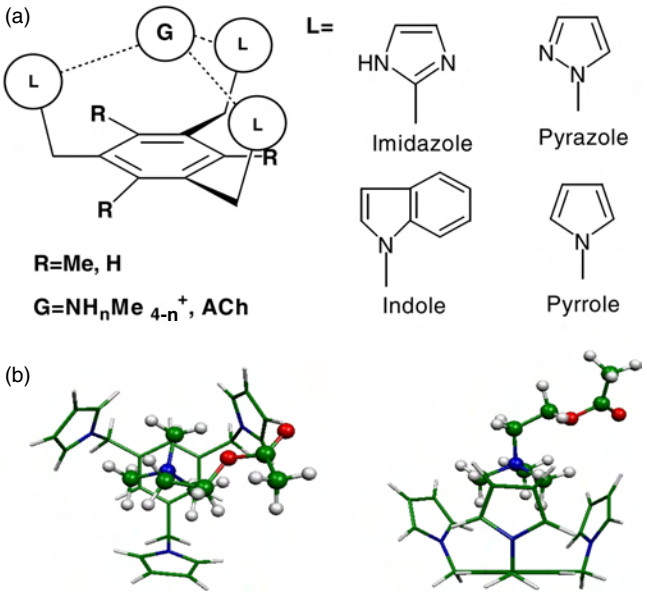


Figure 31. (a) Schematics of tripodal receptors with subunits (1: L = imidazole, 2: L = pyrazole, 3: L = indole, 4: L = pyrrole; a: R = Me, b: R = H). (b) Top and side views of the calculated structures of the complex of 4b with acetylcholine.

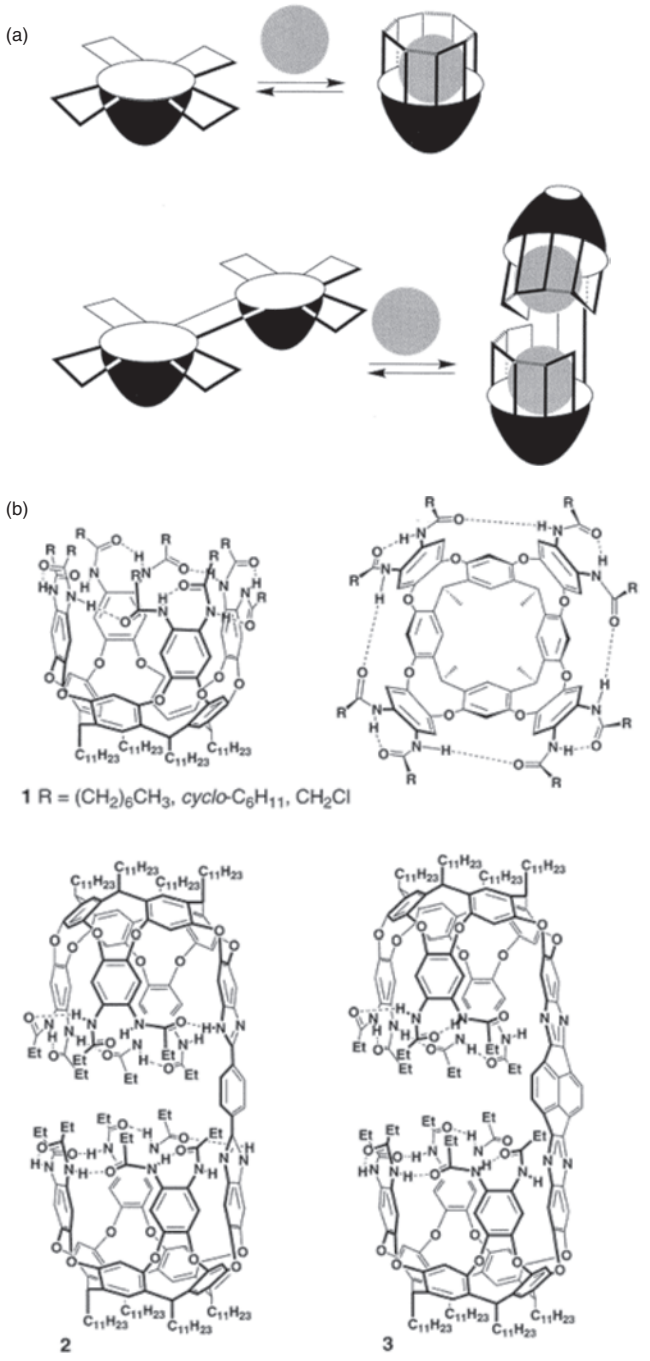


Figure 32. (a) Cartoon representations of self-folding cavitands (top) and self-folding nanoscale containers (bottom) (i.e., unimolecular capsules). (b) Self-folding cavitand 17 (side and top views, only one cycloenantiomer is depicted) and self-folding containers 2 and 3. Reprinted with permission from [479], U. Lücking et al., *J. Am. Chem. Soc.* 122, 8880 (2000). © 2000, American Chemical Society.

accelerated after the appearance of the paper by Hunter and Sanders on the nature of π–π interactions [437]. Though the goal of this paper was to understand the experimental aggregation of porphyrin complexes, the authors defined a set of rules on the use of these π–π interactions, which have been very useful in the design of new molecular entities [15, 481–484]. This exploitation of the

π–π interaction is exemplified in the template directed synthesis of (2)catenane (Fig. 33), which forms continuous π-donor/π-acceptor stacks in the crystal [15]. The photochemical or electrochemical control of the orientations of these π stacks could eventually lead to the fabrication of materials capable of information storage [15]. Another example of the exploitation of these π–π interactions is in the fabrication of a electrochemically controllable nanomechanical molecular system (Fig. 34) [233]. One of the advantages in the harnessing of these π–π interactions, which has already been mentioned, is the fact that the magnitude and orientation can be electrochemically or photochemically controlled. More details of the electrochemical or photochemical control of intermolecular interactions is given in a subsequent section detailing the interaction of quantum species. Another manifestation of the π–π interaction is the stacking of supramolecular porphyrin complexes, which have been conceived as receptors for peptides and nucleotides [485, 486].

In the previous section, we have examined in detail the intermolecular interactions involving π systems. Two issues which have frequently cropped up are self-assembly, self-synthesis, and the electrochemical or photochemical control of the orientations of these π systems. We discuss these issues in the subsequent sections, but before we do so, we discuss the nature of metallic interactions.

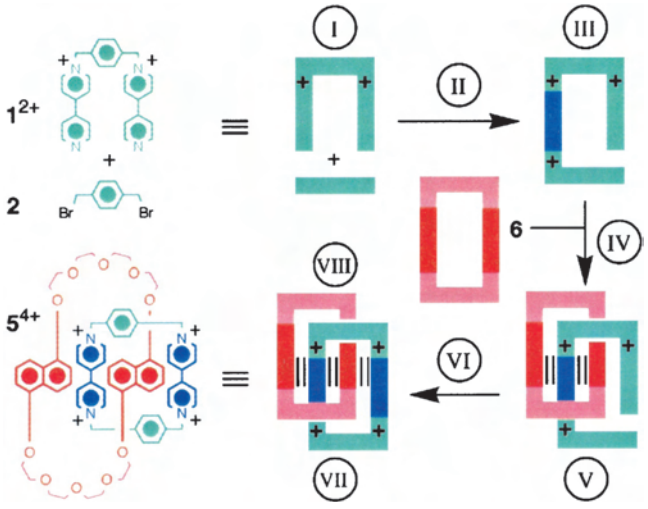


Figure 33. The delicate interplay between covalent, noncovalent, and mechanical bonds in the template-directed synthesis of the (2)catenane 5^{4+}. (I) Dication 1^{2+} does not recognize template 6 until the reaction with dibromide 2 ensues. (II) One covalent bond is formed. (III) A recognition site is switched on in the trication. (IV) Threading is driven by noncovalent bond formation. (V) Molecular recognition is switched on partially in the 1:1 complex. (VI) Another covalent bond is formed. (VII) Yet more molecular recognition is switched on, and a mechanical bond is also formed. (VIII) The weak noncovalent bonding interactions that template the formation of 5^{4+} live on inside the molecule afterward. Reprinted with permission from [15], J. F. Stoddart et al., *Proc. Natl. Acad. Sci. USA* 99, 4797 (2002). © 2002, National Academy of Science.

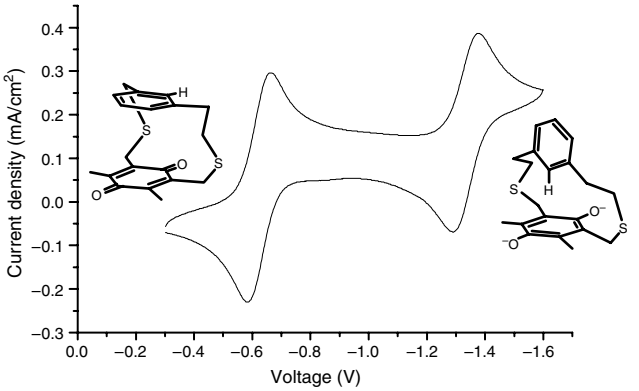


Figure 34. Cyclic voltammogram of 2,11-dithio[4,4]metametaquinocyclophane (1 mM) and tetrabutylammonium perchlorate (0.1 M) in CH_3CN at 25 °C (scan rate 100 mV/s) shows two clear reversible redox reactions, indicating an easy transformation from one structural form to another by simple electrochemical control of the redox reaction. Reprinted with permission from [233], H. G. Kim et al., *Org. Lett.* 4, 3971 (2002). © 2002, American Chemical Society.

4.4. Metallic Interactions

The interaction between metal atoms is generally not discussed in most reviews dealing with conventional molecular recognition. However, in the context of nanorecognition, they are of interest because most metal systems exhibit very unusual properties when their lateral dimensions are decreased to the nano- or subnanoscale [197, 487]. Apart from influencing the structural and energetic characteristics, the interactions between metal atoms also influence the electronic characteristics [8, 9, 197–200, 203–205, 488, 489]. For example, a lead (Pb) subnanowire (lateral dimensions are smaller than a nanometer) becomes a better conductor than a gold (Au) subnanowire by three orders of magnitude. This is in sharp contrast to what is observed at the macroscopic scale, wherein gold is a better conductor than lead by an order of magnitude. Apart from changes in the conductivity, the tensile/cohesive strength of metallic subnanowires is found to be much larger than that observed in bulk. This is because deformations in bulk metallic systems occur through dislocation motions, while structural rearrangements are the norm in materials at the subnanoscale. Recent advances in experimental techniques, especially scanning tunneling microscopy and high-resolution diffraction methods, have led to the unravelling of a number of unusual metallic interactions in the last couple of years [276, 279].

On a different note, conventional wisdom indicates that two interacting open-shell species would exhibit covalent interactions because more bonding orbitals would be filled than antibonding orbitals [487]. While ionic interactions are expected between closed-shell species carrying opposite charges, two closed-shell species can also be expected to interact through a combination of electrostatic, induction, and dispersion interactions. However, it would seem preposterous to expect two closed-shell species carrying similar charges to exhibit an attractive interaction. Indeed, most closed-shell systems carrying similar charges would exhibit a repulsive interaction. However, there is growing evidence that in case of metallic or organometallic species, an entire

family of attractive cation–cation interactions exists between d^8–d^{10}–s^2 systems [490–504]. Though these interactions are much weaker than most covalent or ionic bonds, they are of comparable magnitude to the typical hydrogen bonds [487, 490]. While there is a lot of debate on the origin of this interaction with some reports negating their existence in the case of the lighter elements and some attributing them to packing effects in the solid states, these interactions involving neutral and charged metallic species merit discussion.

4.4.1. Types of Bonding

The interactions in systems containing uncharged metallic species may be classified as either continuous transitions or metallic. In the former interaction, there is a continuous shift from one type of bonding to another type as function of a chemical or physical variable. An interesting feature of this kind of transition is depicted in a covalent/ionic/metallic/van der Waals tetrahedron (Fig. 35) [505], which is based on the Ketelaar–van Arkel triangular diagram.

Though there is no such thing as metallic bonding, metals could be regarded as systems with close-lying HOMOs and LUMOs [506]. Most of the typical properties of metals, conductivity, optical reflectivity, Friedel oscillations, are consequences of these small HOMO–LUMO gaps. When the metallic species are charged, as in case of interactions involving d^{10} gold cations, the interaction seems to be entirely due to dispersion interactions [487]. Additionally it has been noted that these interactions are enhanced by relativistic effects and hence are more pronounced in systems containing heavy metals with large relativisitic effects. Since the goal of most theoretical investigations of large metallic or nonmetallic systems (especially defects in semiconductors, liquid metals, small atomic clusters, phase transitions of solid state systems) is on the physical properties exhibited by the bulk, to our knowledge, there has been no detailed theoretical investigation of the nature of individual metal–metal interactions. Mention should, however, be made of an early report on the experimental observation of a large enhancement of van der Waals forces between small silver particles [507]. The enormous enhancement in the interaction was attributed to a dramatic increase in the dispersion energies.

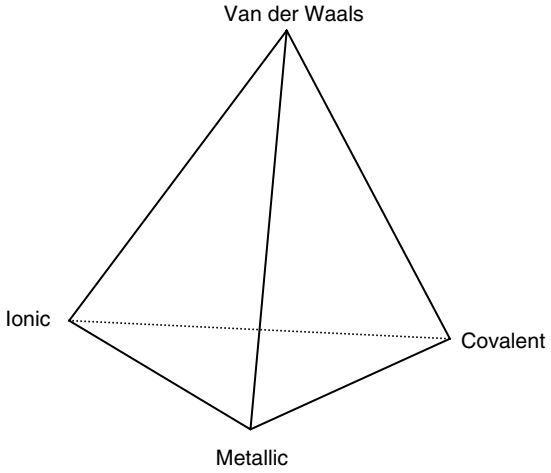


Figure 35. Covalent, ionic, metallic, and van der Waals bond transitions of metal–metal interactions: Laing's tetrahedron model based on the Ketelaar–van Arkel triangular diagram. Reprinted with permission from [505], M. Laing, *Educ. Chem.* 160 (1993). © 1993, Royal Society of Chemistry.

4.4.2. Manifestation

Against this background, we highlight a few metal–metal interactions which have been used to control supramolecular structures and their dimensionality. The foremost among them is aurophilicity, which is the tendency of two closed-shell gold(I) atoms to aggregate at distances shorter than the sum of the van der Waals radii with an interaction energy comparable to that of H-bonds [490]. It was soon realized from both experimental and theoretical studies that these gold–gold interactions are more a consequence of the interaction between the metal centers than being imposed by the ligand architecture [487, 503]. Theoretical efforts in the recent past seem to indicate that analogous phenomena (argentophilicity or cuprophilicity) also occur in interactions of other closed-shell atoms of the same period or even the same group [503]. The experimental studies, however, seem to indicate that most of these Ag(I)–Ag(I) or Cu(I)–Cu(I) interactions are supported by the ligand architecture [508]. This seems to indicate that these interactions are weaker than the corresponding Au(I)–Au(I) interactions.

Among the purely metallic interactions, the recent experimental realization of the icosahedral W@Au_{12} cluster is an interesting example of how theoretical predictions aid the design of novel and viable nanomaterials [204]. The stability of these highly symmetric metallic clusters, containing 12 gold atoms and an encapsulated central impurity atom of the 5*d* elements like tungsten, was theoretically postulated by Pyykkö and Runeberg [205]. Apart from the structural similarity to the analogous C_{60} (Fig. 36), an interesting facet of this metal cluster is that it is stabilized by aurophilic interactions and relativistic effects [204, 205]. The orbital interactions between the W 5*d* and the Au_{12} ligand orbitals lead to the icosahedral bonding h_g HOMO and antibonding h_g^* LUMO, which are separated by a large HOMO–LUMO gap of around 3 eV [204]. These small metallic clusters,

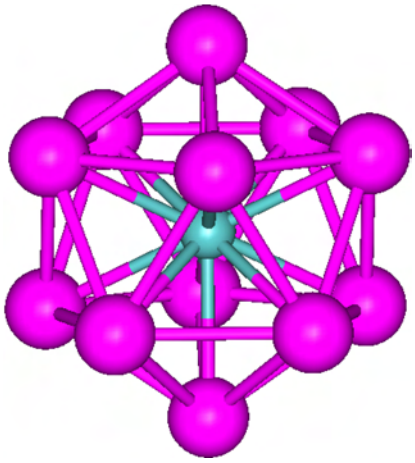


Figure 36. A purely metallic W@Au_{12} cluster, which possesses an I_h symmetry similar to an encapsulated C_{60} fullerene.

especially those of gold, apart from having applications as amplification tags in gene analysis and antigen detection [508–511], have also been implicated for their catalytic roles [512, 513].

This discussion on the role of the encapsulated atom in the stability of the small gold nanocluster leads to an interesting comparison on the role of dopants in the fabrication of nanowires. Given the current interest in the fabrication of nanowires composed of single gold, silver, or copper atoms, there is a need to obtain linear structures. The problem, however, is that these linear structures of gold nanowires are energetically less stable than the corresponding zigzag or bent structures [514, 515]. The alloying of these gold nanowires with carbon or alkali metal atoms seems to circumvent the problem by imbuing more stability on the linear structures [516]. The fact that this linearity can also be obtained by enclosing the single stranded atomic wire in an organic cage can be noted from recent studies [17, 517, 518].

4.5. Interactions Involving Quantum Species

Up to this juncture, most of our discussion on intermolecular interactions involved either molecules, atoms, or stable closed shell cations or anions. In this section, we discuss some unusual types of interactions, which involve quantum species like single electrons, photons, and paramagnetic atoms [15,19, 50, 117, 118, 334, 350, 519–544]. In previous sections, we had mentioned the control of various molecular recognition phenomena by external stimuli: photochemical control involving photons, chemical control involving changes in ion and pH concentration, and electrical control involving electrons. Apart from altering the electronic characteristics, changes in the conformations and orientations of atoms, molecules, or clusters can also be brought about by subtle control of the external stimuli (like in photochemical control, changes in the wavelength of the incident radiation can lead to selective excitation). Though the idea of exerting direct control on molecular recognition phenomena has a biomimetic origin, it has recently been realized that the implications extend to the design of novel nanomaterials. An effective control of molecular recognition phenomena in the nanodomain paves the way to several desirable goals like responsive and intelligent materials, smart sensors, molecular devices, etc. [15, 19, 50, 353].

In a related context, electrical charging has emerged as a central issue in connection with quantal nanostructures like fabricated semiconductor nanodevices like quantum dots [120], gas-phase microsystems like fullerenes [212], large organic molecules [10], molecular clusters, metal, and metal halide clusters [218–220]. Unlike macroscopic systems wherein the number of positive or negative charges that can be added is arbitrary, the number is size controlled at the nanodomain because of quantum instabilities arising from restriction of excess charge in very small spaces. Similar is the case with the interaction of photons with matter at the macroscopic and the nanodomain.

The simplest case of control is through changes in the oxidation states. For example in Figure 37, the β-cyclodextrins do not interact with the redox active guest (viologen) in the most stable +2 oxidation state [61]. However, the reduction of the viologen dramatically enhances the affinity of the violgen toward the host, β-cyclodextrin [61]. This theme of redox control finds application in the design and development of smart sensors, which can be electrochemically controlled [545].

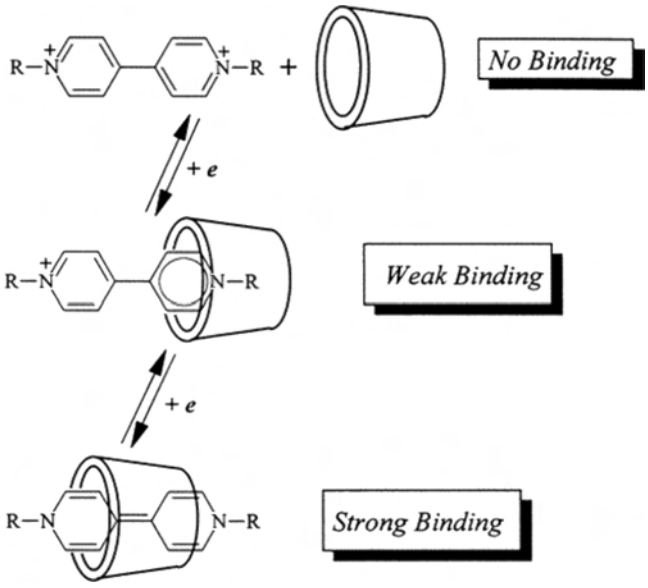


Figure 37. Redox control of the β-cyclodextrin complexation of viologens. Reprinted with permission from [233], A. E. Kaifer, *Acc. Chem. Res.* 32, 62 (1999). © 1999, American Chemical Society.

Electron binding to atoms and molecules has played important roles in physical, chemical, and biological phenomena [521–533]. It has long been known that the presence of an excess electronic charge on biologcal important molecules like DNA and RNA can initiate events which eventually lead to radiation damage, cancer, etc. [521, 522, 546, 547]. In metallurgy, a large number of metals are produced by the electrochemical reduction of the corresponding oxides or sulfides. In both cases, electron binding is known to occur typically by occupation of an empty, or a partially empty, low-energy atomic or molecular orbital, the presence of which is manifested by a positive electron affinity of the system. But in contrast to atoms, a molecular system with a negative electron affinity can still bind an electron by holding the electron in its dipolar electric field. The possibility of such dipole-bound electrons was first proposed by Fermi and Teller in 1947 [548], and many examples have been discovered and studied in molecules and clusters since [115, 197, 526–531, 533]. When the dipole moment of a molecule exceeds 2.5 D, the molecule can bind an electron even when its intrinsic electron affinity is negative [529]. Thus, the acetonitrile anion provides a prototypical case of a ground-state dipole-bound anion molecule [115, 534, 535]. The dipole field that binds the electron can also be supplied by a cluster system even when the individual monomeric unit does not possess a large enough dipole moment. Thus, the water dimer with a dipole moment greater than 2.5 D binds an excess electron, even though the water monomer with a dipole moment of 1.85 D does not [529]. Apart from their potential role in a large number of photochemical and

biological processes [526–531], the issue of the hydrated electron is particularly interesting because with a gradual increase in the size of the water cluster, there is a dramatic change in the nature of the binding of the excess electron [532–535]. While the excess electron in the smaller water clusters is bound to the surface, it is solvated in the larger water clusters, a phenomena which extends itself to bulk water [118]. Given this background, numerous theoretical and experimental attempts have been made to unravel the precise size of the cluster at which there is a transition from a surface bound to a solvated electron [115]. An offshoot of these investigations is that changes in the size and dipole moments can lead to the design and development of effective single electron devices. An example of such a device is the molecular nano-tweezers-like structure obtained for the water hexamer anion (Fig. 38) [117]. Though this structure is not the lowest energy structure of the water hexamer anion, it provides insight into the conformational changes preceding and following the binding of the excess electron.

The formation and stability of the recently discovered multiply charged metal cluster anions is another related and interesting aspect of the binding of excess electrons [218, 549, 550]. In the case of the multiply charged anions, the electrons are bound by a repulsive Coulomb barrier (analogous to that of α-particles in heavy nuclei) [549, 550] and therefore exhibit electronic shell effects arising from magic major shells and subshells associated with ellipsoidal shapes [549, 550]. Given their potential in the isolation and synthesis of size-selected nanoclusters, a lot of effort is currently underway to examine the variation in properties as a result of changes in their sizes and charge states [197, 204, 218, 219].

This discussion about excess electrons brings to fore the related issue of charge transfer to solvent (CTTS) states of hydrated halide anions [531, 542, 551, 552]. The CTTS phenomenon was initially thought to be only of academic interest, but recently it finds importance as a device to trap electrons in polar solvents [552]. The issue of importance herein is that both the energy of the incident radiation and the size of the cluster solvating the halide are crucial to the excitation of the hydrated halide from the ground to the first excited state [552]. While the CTTS states are derived from intermolecular interactions, the corresponding intramolecular charge-transfer transitions in polar solvents find application in the design of fluoroscent molecular reporters [539]. The interaction of electrons or holes with certain kinds of conjugated polymers leads to the formation of singlet and triplet excited states. The excited singlets, which emit light, are indistinguishable from the excited states of photoluminiscence. Conjugated polymers exhibiting this kind of behavior form the basis of electroluminescence and are the impetus for the development of light-emitting devices for display and other purposes [125].

The conformational changes brought about by the interaction of photons with matter, which is the central theme of several biological processes like photosynthesis, vision, etc. [543], are also the basis for the design of molecular switches [12, 541, 542]. Before we end this section, mention should be made of the endohedral $N@C_{60}$ complex (Fig. 39) [106–110, 449]. This endohedral complex, experimentally discovered in 1996, is particularly interesting because electron paramagnetic resonance indicates that the central nitrogen atom is in its highly reactive quartet state [106–110, 449]. However, a large body of both theoretical and experimental evidence seems to indicate that the central nitrogen atom is chemically inert [106–110, 449]. Recent calculations indicate that the complex is entirely stabilized by dispersion interactions, because there is no evidence of any orbital interactions [449]. Against this background, this endohedral $N@C_{60}$ complex is being touted as a viable unit for carrying out quantum computations [553]. Additionally, the ability of C_{60} to enclose highly reactive species has paved the way for a number of interesting applications [110, 449]. It can therefore be noted from this discussion that the absence

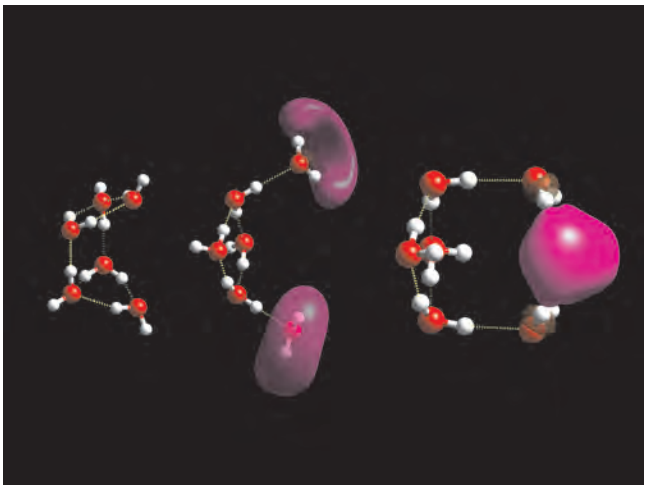


Figure 38. Electron density distributions of HOMO for the initial electron-containing structure (right) and LUMO for a structure of an intermediate electron-released neutral state (middle), and the structure of the final neutral state, upon detaching the excess electron from the initial electron-containing structure in an instant (left). It can be considered to be a prototype of a single-electron device (electron nano-tweezers). Reprinted with permission from [61], K. S. Kim et al., *J. Am. Chem. Soc.* 119, 9329 (1997). © 1997, American Chemical Society.

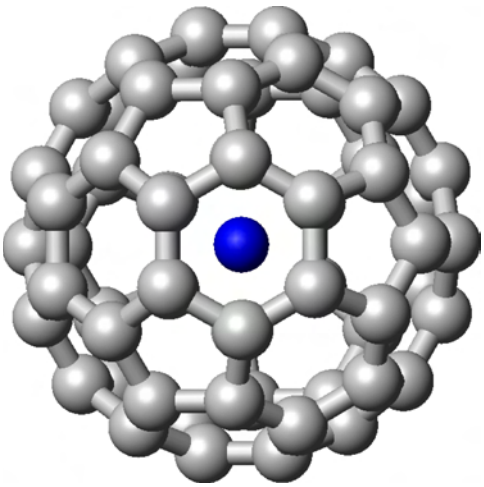


Figure 39. $^4N@C_{60}$. An example of an encapsulated fullerene wherein the highly reactive and paramagnetic nitrogen does not exhibit any orbital interaction with the C_{60} cage. The binding is through dispersion interactions.

of any interaction between two species could also provide useful applications.

5. NANORECOGNITION AND THE QUEST FOR NANOMATERIAL DESIGN

Most of the previous discussion was centered on the principles and various kinds of intermolecular interactions prevailing in nanosystems. Though we examined each interaction in isolation, the properties and structures of most nanosystems are a consequence of the cooperativity and competition of several intermolecular interactions. The importance of controlling the molecular structure and supramolecular order is amply illustrated in the phenyl based polymers–compounds containing phenyl rings linked by a σ-bond [3, 554]. They form an increasingly important class of organic materials because the presence of the phenyl ring aids the construction of a wide range of structures having a large range of active functionality. More importantly, the delocalized π-systems on adjacent rings can overlap to form extended conjugated systems. These π-conjugated organic materials are of considerable importance in materials science in recent years because their optical (e.g., fluorescence) and electrical (e.g., semiconducting) properties make them suitable for use as active components in a variety of electronic devices, including light-emitting diodes, optically pumped lasers, and field-effect transistors [1, 2, 8]. The dimensionality of these polyphenylenes can be exerted in one, two, or even three dimensions. Thus, linear conjugated polymers exhibit conjugation essentially in only one dimension and may be thought of as molecular wires. Disc-like polycyclic aromatic hydrocarbons, which may be regarded as models for graphite, may exhibit two-dimensional conjugation. Three-dimensional conjugation can be obtained from a network containing conjugated segments derived from phenylene-based dendrimers. Since charge transport is a process that is not confined to a single molecule, the stacking of the molecular species to allow 2D and 3D electronic interactions is of crucial importance in giving the high degree of order necessary to obtain an optimal overlap of the π systems, so as to obtain defect-free charge transfer. For example, if we consider 1D polymer chains, then to get good charge transfer between chains, we need to optimize their ability to undergo lamella-type packing. Similarly, with conjugated discs stacking through π interactions to form columns, we can easily envisage efficient 1D charge transport along the axis of the stacking, but to get truly 2D or 3D charge transport, we would need to obtain efficient electronic interactions between the columns.

Against this background on the importance of controlling molecular order and molecular structure, we highlight on how cooperativity and competing effects can be harnessed in nanomaterial design. Finally, we discuss how control of nanorecognition by external stimuli leads to nanodevices and molecular machines.

5.1. Cooperative Effects

One of the definitions of molecular recogntion, proposed by Rebek, is that it arises from intermolecular forces acting on complementary surfaces [38, 66, 479]. As was seen in previous sections, a vast majority of intermolecular interactions are weak and therefore reversible. However, molecules possessing multiple complementary binding sites provide a means to magnify the effect of very weak individual interactions. This effect is more commonly known as cooperativity and is responsible for the physical characteristics of a vast majority of chemical and biological systems. The basis for cooperativity can entirely be explained by the additivity of the corresponding individual free energies and formation constants. Interestingly, entropic arguments need not be invoked to explain this phenomenon. In general, the existence of multiple binding sites in natural or synthetic receptors leads to greater efficiency and selectivity in complexation. Conformational changes during the course of complexation can also lead to the functionally important allosteric systems [56, 57, 66, 191].

One of the widely known manifestations of cooperative effects which is of relevance to this chapter is self-assembly [66, 91–95]. The most interesting aspect of self-assembly is that in a number of cases, the high affinity and selectivity arise from a complementarity of size and shape than from interactions involving charges [57, 66, 191]. The derivatives of glycourils, notably cucurbituril, are one of the most well-known examples of self-assembly [75, 262]. In fact, the large number of fascinating molecular architectures (Fig. 40) formed from cucurbiturils epitomize the salient features of self-assembly [75, 83] and also lead to the concept of self-assembling capsules.

We can define self-assembling capsules as receptors with enclosed cavities that are formed by the reversible noncovalent interaction of two or more, not necessarily identical, subunits. The aggregate so formed should have a well-defined structure in solution and be capable of binding behavior that none of its individual components display alone. This definition naturally places an emphasis on three-dimensional assemblies. It specifically excludes indefinite arrays and aggregates that are not designed to act as receptors such as micelles, liquid crystals, monolayers, metal–ligand structures, and assemblies that exist only in the solid state. As such, self-assembling capsules represent a subset of self-assembling systems. Reversibly formed capsules, held together by intermolecular forces, provide one means of accomplishing recognition under equilibrium conditions (Fig. 32). This approach also gives rise to larger molecular containers that can encapsulate increasingly larger guests and incorporate the properties required for, say, catalysis [84].

5.2. Competing Effects

A number of studies on H-bonding in organic molecular assemblies have revealed that certain classes of functional groups always form H-bonds when complementary donors or acceptors are available, while other classes of functional groups only occasionally participate in H-bonding. The discussion on the concept of best donor and acceptor forming a H-bond provides a cogent tool for controlling the composition and structure of molecular assemblies. Thus, novel molecular assemblies can be designed using complementary sets of strong and weak H-bonds [57, 331, 555]. It is interesting to note that this competition between H-bonds in solid state mirrors those found in solution (or in the gas phase) [62, 87, 555].

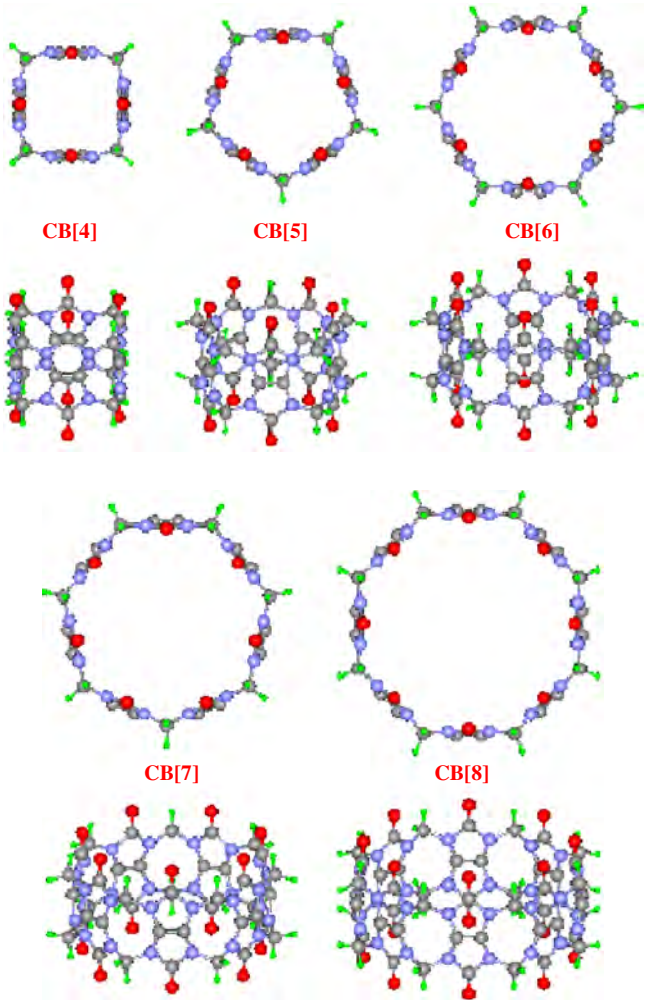


Figure 40. Self-assembly. Examples of a wide variety of molecular architectures obtained from different sizes of cucurbiturils. Adapted in part with permission from [61], J. Kim et al., *J. Am. Chem. Soc.* 122, 540 (2000). © 2000, American Chemical Society.

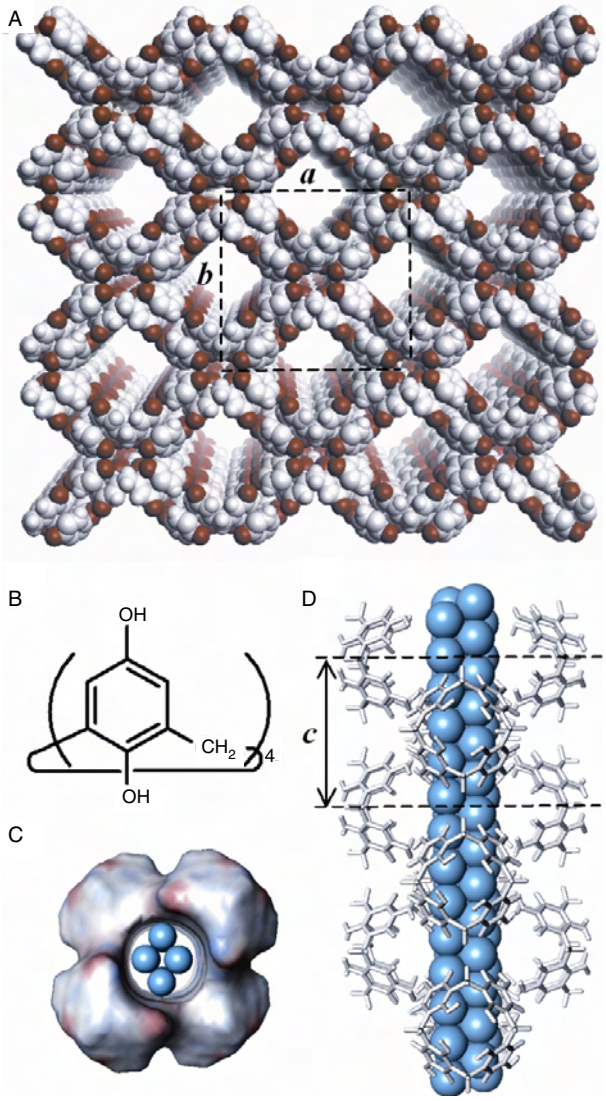


Figure 41. Fabrication of ultrathin silver nanowires from self-assembled CHQ nanotubes. (A) A schematic view of CHQ nanotube arrays based on the X-ray analysis of the crystal reveals that the pores of 0.8 × 0.8 nm^2 (with van der Waals volume excluded) are separated by 1.7 nm from the neighboring ones. O atoms are shown in red. (B) CHQ monomer. (C and D) Top and side views, respectively, of a silver nanowire inside a CHQ nanotubetemplate. The solvent-accessible surface model in (C) and the stick model in (D) show schematic views of the CHQ nanotubes. The red color of the surface represents the negative electrostatic potential on oxygen atoms. The space-filled models (blue) represent the silver nanowires. Reprinted with permission from [17], B. H. Hong et al., *Science* 294, 348 (2001). © 2001, American Assocation for Advancement of Science.

While there is a marked preference for the formation of strong H-bonds, it is of interest to note that weak H-bonds behave as steering interactions. Thus, Desiraju demonstrated that H-bonds, other than those taking part in strong H-bonds, play a crucial role in determining crystal packing [431, 555].

A stellar example of the use of competing effects in the the design of nanoarchitectures can be noted from a recent study, wherein the 1D short H-bonds and the displaced π–π stacking interactions in the calix[4]hydroquinone (CHQ) nanotube bundles compete in the assembling process (Fig. 1) [16, 18]. Incidentally, these CHQ nanotubes were used as templates in the fabrication of ultrathin silver nanowires (Fig. 41) [17, 31].

5.3. Nanodevices and Molecular Machines

As is the case with most issues discussed in this chapter, the concept of a nanodevice, or for that matter, a molecular machine, has its inspiration in nature [11–13, 15, 556]. One of the most spectacular examples of a nanodevice incorporating different aspects of nanorecognition and control is the case of the muscle linear and ATP-ase rotary motors [557].

A variety of molecular and supramolecular systems have been designed in the recent past, in which conformational changes, movements, occur as a response to external chemical, electrical, or photochemical stimuli [11, 12, 15, 233, 558, 559]. One of the advantages of the use of nanorecognition in this process is that it offers a (bottom-up) approach, which

is unlike the (top-down) approach favored by physicists and engineers. The limitations of the latter approach can already be noted in several systems.

6. CONCLUSIONS

In this review, we have discussed the principles underlying nanorecognition and various manifestations of intermolecular interactions in nanosystems. Some of the salient conclusions which can be drawn are outlined below.

(i) Nanorecognition, in addition to conventional molecular recognition, also incorporates the interaction of quantum species like electrons, photons, etc.
(ii) Given the small size of the systems involved, energetic considerations seem to be the more dominant indicator of stability in nanorecognition.
(iii) Quantum effects have to be taken into account to obtain realistic descriptions of most intermolecular interactions in nanorecognition.
(iv) Though vast strides have been made in obtaining quantitative estimates of the interaction energies of most nanosystems, the necessity of taking into account the quantum effects implies that the investigations can be carried out for very small and simple systems.

As can be seen, an understanding of nanorecognition would (i) aid the development of novel nanosystems with potential applications in pharmaceuticals, optics, optoelectronics, information storage, sensors, biotechnology, nanomedicine, nanofluidics, nanoelectronics, (ii) provide information on the methods needed to harness these nanosystems as machines and devices, and (iii) provide an increased understanding of the science behind these nanoscale processes.

GLOSSARY

Competing interactions When several kinds of intermolecular interaction prevail in a system, the magnitudes of some of them are enhanced and the others depleted. In larger systems, the interactions which are depleted are generally not considered. However, in nano dimensions, these very weak interactions have subtle effects which cannot be neglected.

Cooperativity The magnification of the effects (structural, energetic, and thermodynamic), when several intermolecular interactions prevail simultanaously in a system. In general, cooperativity is enhanced when an initial interaction makes the system more conducive for the next attractive recognition process.

Dispersion energy The attractive energy arising from the interaction of the fluctuating charge distributions of the interacting systems as induced-dipole–induced-dipole interactions. These fluctuations result from the correlation of the motion of the electrons in the system and thereby favor lower energy configurations over higher energy ones.

Electron correlation Electron correlation represents the effects which are manifested due to the instantaneous interaction between electrons.

Electron density The probability of finding any electron of a molecule or an intermolecular complex at a given point in space. The electron density is obtained by integrating the square of the modulus of an electronic wave function.

Electronic wave function The wave function Ψ is a function of the electron and nuclear positions and is based on the description of an electron as a wave.

Electrostatic energy The energy arising from the interaction of the static charge distributions of the interacting systems. They can either be attractive or repulsive.

Exchange–repulsion The repulsion due to the Pauli exclusion principle, which arises from the overlap of the occupied molecular orbitals of the interacting systems.

Hartree–Fock The Schrödinger equation is solved by assuming that the motions of particles constituting the system are independent of each other and a given particle interacts with the averaged charged distribution of other particles in the system. The multielectron wave function is described by an approximation of the unsymmetrized product of one-electron functions (orbitals).

Hydrogen bond An attractive interaction mediated through a hydrogen atom. A typical example is an interaction of the type **D**–H$\cdots$**A**, wherein **D** is a proton donor and **A** is a proton acceptor. Other examples include an interaction of the type **A**–H$\cdots$**A**.

Hydrophobicity The phenomenon of clustering of nonpolar molecules or groups in the presence of water. The driving force for this clustering is not due to the strong attraction between the nonpolar groups, but because of the high affinity of water molecules to form hydrogen bonds between themselves.

Induction energy The attractive energy arising from the interaction of the static charge distribution with the induced charge distribution of the other.

Interaction energy The sum of all the attractive and repulsive energies between two interacting systems. The more negative the interaction energy, the stronger the interaction.

Molecular recognition The processes leading to the binding and selection of substrate(s) by a receptor and culminating in some function. These processes are a result of a well-defined pattern of intermolecular interactions.

Nanorecognition Nanorecognition is synonymous with molecular recognition in most cases. However, the small size of the systems involved implies that quantum effects govern the nature of intermolecular interactions. In particular, the interactions involving electrons and photons are also included in nanorecognition.

Preorganized Molecules, in particular host molecules, which do not undergo significant conformational reorganization upon binding to a guest are said to be preorganized.

Quantum effects They represent those effects which need to take into account the dual nature (particle and wave) of the systems involved to obtain a reasonable description of their physical characteristics. Though these effects can be neglected when the systems are large, they become very important when the system under consideration is extremely small.

Relativistic effects These effects owe their existence to the very fast motion of electrons closest to the nuclei of heavy atoms.

Self-assembly The spontaneous organization of molecules under thermodynamic equilibrium conditions into structurally well-defined and rather stable arrangements through a number of noncovalent interactions.

Self-synthesis What distinguishes self-assembly from self-synthesis is that in the former, the identity of the molecules or molecular template is maintained but in the latter, an entirely new chemical system, bearing no resemblance to the original molecular template, is formed.

Supermolecule Supermolecule denotes singular or polymolecular associations which are well-defined, discrete oligomolecular species that result from the intermolecular association of a few components.

Supramolecule Supramolecules result from the spontaneous association of a large undefined number of components into a specific phase having more or less well-defined microscopic organization and macroscopic characteristics depending on its nature.

Van der Waals forces One of the older classifications of weak intermolecular forces, which represent the weak attractive and repulsive forces prevailing in an ensemble of molecules. The chemical identities of the constituent molecules are not altered as a result of van der Waals interaction. In physical terms, they represent the forces arising from not only dipole–dipole, dipole–induced-dipole or dispersion interactions but also exchange–repulsion.

ACKNOWLEDGMENTS

This work was supported by the Korean Ministry of Science and Technology under the Creative Research Initiatives Program. We thank Professor Michael Ward for providing us the images needed for making Figure 22, and Professor Lai-Sheng Wang for providing the coordinates of $W@Au_{12}$.

REFERENCES

1. A. P. Alivisatos, P. F. Barbara, A. W. Castleman, J. Chang, D. A. Dixon, M. L. Klein, G. L McLendon, J. S. Miller, M. A. Ratner, P. J. Rossky, S. I. Stupp, and M. E. Thompson, *Adv. Mater.* 10, 1297 (1998).
2. J. R. Heath, *Acc. Chem. Res.* 32, 388 (1999).
3. S. Becker and K. Müllen, in "Stimulating Concepts in Chemistry" (M. Shibasaki, J. F. Stoddart, and F. Vögtle, Eds.), p. 317. Wiley–VCH, Weinheim, 2000.
4. S. H. Gellman, *Chem. Rev.* 97, 1231 (1997).
5. F. Vögtle, "Supramolecular Chemistry: An Introduction." Wiley, Chichester, UK, 1993.
6. J.-M. Lehn, "Supramolecular Chemistry." VCH, Weinheim, 1995.
7. "Comprehensive Supramolecular Chemistry," Vols. 1–12. (J.-M. Lehn, Ed.). Pergamon, Oxford, 1996.
8. "Nanostructure Physics and Fabrication" (M. A. Reed and W. P. Kirk, Eds.). Academic, San Diego, 1989.
9. "Nanostructures and Mesoscopic Systems" (W. P. Kirk and M. A. Reed, Eds.). Academic, San Diego, 1991.
10. J. M. Tour, *Chem. Rev.* 96, 537 (1996).
11. R. Ballardini, V. Balzani, A. Credi, M. T. Gandolfi, and M. Venturi, *Acc. Chem. Res.* 34, 445 (2001).
12. B. L. Feringa, *Acc. Chem. Res.* 34, 504 (2001).
13. J. P. Spatz, *Angew. Chem. Int. Ed.* 18, 3359 (2002).
14. D. G. Castner and B. D. Ratner, *Surf. Sci.* 500, 28 (2002).
15. J. F. Stoddart and H.-R. Tseng, *Proc. Natl. Acad. Sci. USA* 99, 4797 (2002).
16. B. H. Hong, C.-W. Lee, J. Y. Lee, J. C. Kim, and K. S. Kim, *J. Am. Chem. Soc.* 123, 10748 (2001).
17. B. H. Hong, S. C. Bae, C.-W. Lee, S. Jeong, and K. S. Kim, *Science* 294, 348 (2001).
18. K. S. Kim, S. B. Suh, J. C. Kim, B. H. Hong, E. C. Lee, S. Yun, P. Tarakeshwar, J. Y. Lee, Y. Kim, H. Ihm, H. G. Kim, J. W. Lee, J. K. Kim, H. M. Lee, D. Kim, C. Cui, S. J. Youn, H. Y. Chung, H. S. Choi, C.-W. Lee, S. J. Cho, S. Jeong, and J.-H. Cho, *J. Am. Chem. Soc.* 124, 14268 (2002).
19. Y. Xia, J. A. Rogers, K. E. Paul, and G. M. Whitesides, *Chem. Rev.* 99, 1823 (1999).
20. C. Sanchez, G. J. de A. A. Soler-Illia, F. Ribot, T. Lalot, C. R. Mayer, and V. Cabuil, *Chem. Mater.* 13, 3061 (2001).
21. G. J. de A. A. Soler-Illia, C. Sanchez, B. Lebeau, and J. Patarin, *Chem. Rev.* 102, 4093 (2002).
22. F. Zeng, S. C. Zimmerman, S. V. Kolotuchin, D. E. C. Reichert, and Y. Ma, *Tetrahedron* 58, 825 (2002).
23. A. K. Boal, M. Gray, F. Ilhan, G. M. Clavier, L. Kapitzky, and V. M, Rotello, *Tetrahedron* 58, 765 (2002).
24. S. F. Bent, *Surf. Sci.* 500, 879 (2002).
25. B. D. Ratner and H. Shi, *Curr. Opin. Solid State. Mater. Sci.* 4, 395 (1999).
26. C. M. Niemeyer, *Rev. Mol. Biotech.* 82, 47 (2001).
27. H. Rohrer, *Microelectron. Eng.* 41/42, 31 (1998).
28. M. A. Ratner, *Mater. Today* 5, 20 (2002).
29. M. A. Reed, *Proc. IEEE* 87, 652 (1999).
30. M. L. Cohen, *Mater. Sci. Eng. C* 15, 1 (2001).
31. K. S. Kim, *Curr. Appl. Phys.* 2, 65 (2002).
32. F. Diederich, *Chimia* 55, 821 (2001).
33. L. M. Greig and D. Philip, *Chem. Soc. Rev.* 30, 287 (2001).
34. M. Shimomura and T. Sawadaishi, *Curr. Opin. Coll. Int. Sci.* 6, 11 (2001).
35. N. C. Seeman, *Curr. Opin. Struct. Biol.* 6, 519 (1996).
36. J. S. Brodbelt, *Int. J. Mass Spectrom.* 200, 57 (2000).
37. C. Schalley, *Mass. Spectrom. Rev.* 20, 253 (2002).
38. F. Hof and J. Rebek, *Proc. Natl. Acad. Sci. USA* 99, 4775 (2002).
39. V. M. Rotello, in "Electron Transfer in Chemistry" (V. Balzani, Ed.), p. 68. Wiley–VCH, Weinheim, 2002.
40. M. D. Hollingsworth, *Science* 295, 2410 (2002).
41. A. Gavezzotti, *Synlett* 2, 201 (2002).
42. N. S. Sampson, M. Mrksich, and C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA* 98, 12870 (2001).
43. E. T. Kool, *Ann. Rev. Biophys. Biomol. Str.* 30, 1 (2001).
44. J. K. M. Sanders, *Pure Appl. Chem.* 72, 2265 (2000).
45. P. D. Beer and P. A. Gale, *Angew. Chem. Int. Ed. Engl.* 40, 486 (2001).
46. B. P. Hay and R. D. Hancock, *Coord. Chem. Rev.* 212, 61 (2001).
47. M. J. Kirsche and J.-M. Lehn, *Struct. Bond.* 96, 3 (2000).
48. P. B. Armentrout, *Int. J. Mass Spectrom.* 193, 227 (1999).
49. C. A. Schalley, *Int. J. Mass Spectrom.* 194, 11 (2000).
50. J. J. Lavigne and E. V. Anslyn, *Angew. Chem. Int. Ed. Engl.* 40, 3118 (2001).
51. P. B. Dervan, *Bioorg. Med. Chem.* 9, 2215 (2001).
52. M. M. G. Antonisse and D. N. Reinhoudt, *Chem. Commun.* 443 (1998).
53. N. Zacharias and D. A. Dougherty, *Trends Pharmacol. Sci.* 23, 281 (2002).
54. G. W. Gokel, S. L. D. Wall, and E. S. Meadows, *Eur. J. Org. Chem.* 2967 (2000).
55. M. D. Ward and A. M. Pivovar, *Curr. Opin. Solid State Mater. Sci.* 4, 581 (1999).
56. H.-J. Schneider, *Angew. Chem. Int. Ed. Engl.* 30, 1417 (1991).

57. D. N. Chin, J. A. Zerkowski, J. C. MacDonald, and G. M. Whitesides, in "Organized Molecular Assemblies in the Solid State" (J. K. Whitesell, Ed.), Ch. 5. Wiley, New York, 1999.
58. S. R. Seidel and P. J. Stang, *Acc. Chem. Res.* 35, 972 (2002).
59. V. G. Machado, P. N. W. Baxter, and J.-M. Lehn, *J. Braz. Chem. Soc.* 12, 431 (2001).
60. J.-H. Fournier, T. Maris, J. D. Wuest, W. Guo, and E. Galoppini, *J. Am. Chem. Soc.* 125, 1002 (2003).
61. A. E. Kaifer, *Acc. Chem. Res.* 32, 62 (1999).
62. G. R. Desiraju, *J. Chem. Soc., Dalton Trans.* 3745 (2000).
63. C. V. K. Sharma, *Cryst. Growth Des.* 2, 465 (2002).
64. A. Ciferri, *Progr. Polym. Sci.* 20, 1081 (1995).
65. M. Sastry, M. Rao, and K. N. Ganesh, *Acc. Chem. Res.* 35, 847 (2002).
66. M. M. Conn and J. Rebek, *Chem. Rev.* 97, 1647 (1997).
67. H.-J. Böhm and G. Klebe, *Angew. Chem. Int. Ed. Engl.* 35, 2588 (1996).
68. A. W. Bosman, R. Vestberg, A. Heumann, J. M. J. Frechet, and C. J. Hawker, *J. Am. Chem. Soc.* 125, 715 (2003).
69. M. Nishio, Y. Umezawa, M. Hirota, and Y. Takeuchi, *Tetrahedron* 51, 8665 (1995).
70. S. J. Cho, H. S. Hwang, J. M. Park, K. S. Oh, and K. S. Kim, *J. Am. Chem. Soc.* 118, 485 (1996).
71. S. B. Suh, C. Cui, H. S. Son, J. S. U, Y. Won, and K. S. Kim, *J. Phys. Chem. B* 106, 2061 (2002).
72. C. Cui and K. S. Kim, *J. Phys. Chem. B* 103, 2751 (1999).
73. K. S. Kim, C. Cui, and S. J. Cho, *J. Phys. Chem. B* 102, 461 (1998).
74. H. S. Choi, S. B. Suh, S. J. Cho, and K. S. Kim, *Proc. Natl. Acad. Sci. USA* 95, 12094 (1998).
75. K. Kim, *Chem. Soc. Rev.* 31, 96 (2002).
76. H.-J. Kim, W. S. Jeon, Y. H. Ko, and K. S. Kim, *Proc. Natl. Acad. Sci. USA* 99, 5007 (2002).
77. J. Chin, J. Oh, S. Y. Jon, S. H. Park, C. Walsdorff, B. Stranix, A. Ghoussoub, S. J. Lee, H. J. Chung, S.-M. Park, and K. Kim, *J. Am. Chem. Soc.* 124, 5374 (2002).
78. S. Y. Jon, J. Kim, M. Kim, S.-H. Park, W. S. Jeon, J. Heo, and K. Kim, *Angew. Chem. Int. Ed. Engl.* 40, 2116 (2001).
79. J. Chin, C. Walsdorff, B. Stranix, J. Oh, H. J. Chung, S.-M. Park, and K. Kim, *Angew. Chem. Int. Ed. Engl.* 38, 2756 (1999).
80. D. Whang, K.-M. Park, J. Heo, P. Ashton, and K. Kim, *J. Am. Chem. Soc.* 120, 4899 (1998).
81. D. Whang, J. Heo, J. H. Park, and K. Kim, *Angew. Chem. Int. Ed. Engl.* 37, 78 (1998).
82. Y.-M. Jeon, J. Kim, D. Whang, and K. Kim, *J. Am. Chem. Soc.* 118, 9790 (1996).
83. J. Kim I.-S. Jung, S.-Y. Kim, E. Lee, J.-K. Kang, S. Sakamoto, K. Yamaguchi, and K. Kim, *J. Am. Chem. Soc.* 122, 540 (2000).
84. J.-S. Seo, D. Whang, H. Lee, S. I. Jun, J. Oh, Y. J. Jeon, and K. S. Kim, *Nature* 404, 982 (2000).
85. G. Wipff, P. A. Kollman, and J. M. Lehn, *J. Mol. Str. THEOCHEM* 93, 153 (1983).
86. H.-J. Schneider and A. Yatsimirski, "Principles and Methods in Supramolecular Chemistry." Wiley, New York, 2000.
87. G. R. Desiraju and T. Steiner, "The Weak Hydrogen Bond in Structural Chemistry and Biology." Oxford Univ. Press, Oxford, 1999.
88. M. Nishio, M. Hirota, and Y. Umezawa, "The CH–π Interaction: Evidence, Nature, and Consequences." Wiley–VCH, New York, 1998.
89. P. D. Beer, P. A. Gale, and D. K. Smith, "Supramolecular Chemistry." Oxford Univ. Press, Oxford, 1999.
90. J. W. Steed and J. L. Atwood, "Supramolecular Chemistry." Wiley, New York, 2000.
91. H. Kuhn, "Organized Monolayers: Components of Artificial Supramolecular Machines." Wiley, New York, 2002.
92. L. F. Lindoy and I. M. Atkinson, "Self-Assembly in Supramolecular Systems." Royal Society of Chemistry, London, 2000.
93. A. Ciferri, "Supramolecular Polymers." Dekker, New York, 2000.
94. H. Dodziuk, "Introduction to Supramolecular Chemistry." Kluwer, Dordrecht, 2002.
95. H. Kuhn and H.-D. Försterling, "Principles of Physical Chemistry: Understanding Molecules, Molecular Assemblies, Supramolecular Machines." Wiley, New York, 1999.
96. C. Santori, D. Fattal, J. Vuckovic, G. S. Solomon, and Y. Yamamoto, *Nature* 419, 594 (2002).
97. C. Joachim, J. K. Gimzewski, and A. Aviram, *Nature* 408, 541 (2000).
98. S. Komiyama, O. Astafiev, V. Antonov, T. Kutsuwa, and H. Hirai, *Nature* 403, 405 (2000).
99. H. Park, J. Park, A. K. L. Lim, E. H. Anderson, A. P. Alivisatos, and P. L. McEuen, *Nature* 407, 57 (2000).
100. J. Park, A. N. Pasupath, J. Goldsmith, C. Chang, Y. Yaish, J. R. Petta, M. Rinkoski, J. P. Sethna, H. D. Abruñas, P. L. McEuen, and D. C. Ralph, *Nature* 417, 722 (2002).
101. C. Dekker, *Phys. Today* 52, 22 (1999).
102. L. Kouwenhoven and L. Glazman, *Phys. World* 14, 33 (2001).
103. N. Kimizuka, *Adv. Mater.* 12, 1461 (2000).
104. A. I. Yanson, G. R. Bollinger, H. E. van den Brom, N. Agraït, and J. M. van Ruitenbeek, *Nature* 395, 783 (1998).
105. A. Chu, J. Cook, R. J. R. Heesom, J. L. Hutchison, M. L. H. Green, and J. Sloan, *Chem. Mater.* 8, 2751 (1996).
106. T. Almeida Murphy, T. Pawlik, A. Weidinger, M. Höhne, R. Alcala, and J. M. Spaeth, *Phys. Rev. Lett.* 77, 1075 (1996).
107. H. Mauser, N. J. R. van Eikema Hommes, T. Clark, A. Hirsch, B. Pietzak, A. Weidinger, and L. Dunsch, *Angew. Chem. Int. Ed. Engl.* 36, 2835 (1997).
108. A. Weidinger, M. Waiblinger, B. Pietzak, and T. Almeida Murphy, *Appl. Phys. A* 66, 287 (1998).
109. B. Pietzak, M. Waiblinger, T. Almeida Murphy, A. Weidinger, M. Höhne, E. Dietel, and A. Hirsch, *Chem. Phys. Lett.* 279, 259 (1997).
110. E. Dietel, A. Hirsch, B. Pietzak, M. Waiblinger, K. Lips, A. Weidinger, A. Gruss, and K. P. Dinse, *J. Am. Chem. Soc.* 121, 2432 (1999).
111. K.-P. Dinse, *Phys. Chem. Chem. Phys.* 4, 5442 (2002).
112. R. L. Whetten, M. N. Shafigullin, J. T. Khoury, T. G. Schaaff, I. Vezmar, M. M. Alvarez, and A. Wilkinson, *Acc. Chem. Res.* 32, 397 (1999).
113. A. Caneschi, D. Gatteschi, N. Lalioti, C. Sangregorio, and R. Sessoli, *J. Chem. Soc., Dalton Trans.* 3907 (2000).
114. A. D. Miller, I. Bezel, K. J. Gaffney, S. Garrett-Roe, S. H. Liu, P. Szymanski, and C. B. Harris, *Science* 297, 1163 (2002).
115. C. Desfrancois and J. P. Schermann, *Chem. Soc. Rev.* 31, 269 (2002).
116. M. Mitsui, A. Nakajima, and K. Kaya, *J. Chem. Phys.* 117, 9740 (2002).
117. K. S. Kim, S. Lee, J. Kim, and J. Y. Lee, *J. Am. Chem. Soc.* 119, 9329 (1997).
118. K. S. Kim, I. Park, S. Lee, K. Cho, J. Y. Lee, J. Kim, and J. D. Joannopoulos, *Phys. Rev. Lett.* 76, 956 (1996).
119. R. L. Carroll and C. B. Gorman, *Angew. Chem. Int. Ed. Engl.* 41, 4378 (2002).
120. C. G. Rocha, T. G. Dargam, and A. T. Latge, *Phys. Status. Solid. B* 232, 37 (2002).
121. J. M. Tour, A. M. Rawlett, M. Kozaki, Y. Yao, R. C. Jagessar, S. M. Dirk, D. W. Price, M. A. Reed, C.-W. Zhou, J. Chen, W. Wang, and I. Campbell, *Chem. Eur. J.* 7, 5118 (2001).
122. E. Coronado, M. Clemente-Léon, J. R. Galán-Mascarós, C. Giménez-Saiz, C. J. Gómez-García, and E. Martínez-Ferrero, *J. Chem. Soc., Dalton Trans.* 3955 (2000).
123. Y. Cai and S. L. Bernasek, *J. Am. Chem. Soc.* 125, 1655 (2003).
124. J. Zhang, Q. Chi, A. M. Kuznetsov, A. G. Hansen, H. Wackerbarth, H. E. M. Christensen, J. E. T. Anderson, and J. Ulstrup, *J. Phys. Chem. B* 106, 1131 (2002).

125. A. Kraft, A. C. Grimsdale, and A. B. Holmes, *Angew. Chem. Int. Ed. Engl.* 37, 402 (1998).
126. S. Horiuchi, Y. Okimoto, R. Kumai, and Y. Tokura, *Science* 299, 229 (2003).
127. M. Shim, N. W. S. Kam, R. J. Chen, Y. Li, and H. Dai, *Nano Lett.* 2, 285 (2002).
128. W. Cai, Y. Zhang, J. Jia, and L. Zhang, *Appl. Phys. Lett.* 73, 2709 (1998).
129. N. Sata, K. Eberman, and J. Maier, *Nature* 408, 946 (2000).
130. K. A. Williams, P. T. M. Veenhuizen, B. G. de la Torre, R. Erjita, and C. Deeker, *Nature* 420, 761 (2002).
131. C. M. Lieber, *Solid State Commun.* 107, 607 (1998).
132. C. Zhou, M. R. Deshpande, M. A. Reed, L. Jones II, and J. M. Tour, *Appl. Phys. Lett.* 71, 611 (1997).
133. M. E. McHenry and D. E. Laughlin, *Acta Mater.* 48, 223 (2000).
134. J. Huskens, M. A. Deij, and D. N. Reinhoudt, *Angew. Chem. Int. Ed. Engl.* 41, 4467 (2002).
135. P. Avouris, *Acc. Chem. Res.* 35, 1026 (2002).
136. A. J. Kirby, *Angew. Chem. Int. Ed. Engl.* 35, 707 (1996).
137. T. C. Bruice and S. J. Benkovic, *Biochemistry* 39, 6267 (2000).
138. W. R. Cannon and S. J. Benkovic, *J. Biol. Chem.* 273, 26257 (1998).
139. A. Warshel, *J. Biol. Chem.* 273, 27035 (1998).
140. B. Nidetzky, C. Eis, and M. Albert, *Biochem. J.* 351, 649 (2000).
141. R. Wolfenden, *Bioorgan. Med. Chem.* 7, 647 (1999).
142. W. Neuhauser, D. Haltrich, K. D. Kulbe, and B. Nidetzky, *Biochemistry* 37, 1116 (1998).
143. J. Steyaert, *Eur. J. Biochem.* 247, 1 (1997).
144. J. F. Sebastian, G. Q. Liang, A. Jabarin, K. Thomas, and H. B. Wu, *Bioorg. Chem.* 24, 290 (1996).
145. M. N. Namchuk and S. G. Withers, *Biochemistry* 34, 16194 (1995).
146. K. S. Kim, K. S. Oh, and J. Y. Lee, *Proc. Natl. Acad. Sci. USA* 97, 6373 (2000).
147. J. K. Park, S. J. Cho, S. Lee, K. S. Kim, and D. H. Kim, *J. Biomol. Struct. Dyn.* 12, 1033 (1995).
148. K. S. Oh, S.-S. Cha, D.-H. Kim, H.-S. Cho, N.-M. Ha, G. Choi, J. Y. Lee, P. Tarakeshwar, H. S. Son, K. Y. Choi, B.-H. Oh, and K. S. Kim, *Biochemistry* 39, 13891 (2000).
149. K. A. Eidne, K. M. Kroeger, and A. C. Hanyaloglu, *Trends Endocrinol. Metab.* 13, 415 (2002).
150. N. Vasudevan, S. Ogawa, and D. Pfaff, *Physiol. Rev.* 82, 923 (2002).
151. A. Bisello, A. E. Adams, D. F. Mierke, M. Pellegrini, M. Rosenblatt, L. J. Suva, and M. Chorev, *J. Biol. Chem.* 273, 22498 (1998).
152. P. Wallimann, T. Marti, A. Fürer and F. Diederich, *Chem. Rev.* 97, 1567 (1997).
153. X. L. Jiang, M. Dreano, D. R. Buckler, S. Cheng, A. Ythier, H. Wu, W. A. Hendrickson, and N. ElTayar, *Structure* 3, 1341 (1995).
154. R. Jimenez, G. Salazar, K. K. Baldridge, and F. E. Romesberg, *Proc. Natl. Acad. Sci. USA* 100, 92 (2003).
155. W. Z. Xie, D. Wang, H. W. Du, and J. Cheng, *Progr. Biochem. Biophys.* 29, 311 (2002).
156. D. P. Allison, P. Hinterdorfer, and W. H. Han, *Curr. Opin. Biotech.* 13, 47 (2002).
157. U. Dammer, M. Hegner, D. Anselmetti, P. Wagner, M. Dreier, W. Huber, and H. J. Guntherodt, *Biophys. J.* 70, 2437 (1996).
158. G. J. Wedemayer, P. A. Patten, L. H. Wang, P. G. Schulz, and R. C. Stevens, *Science* 276, 1665 (1997).
159. P. A. Patten, N. S. Gray, P. L. Yang, C. B. Marks, G. J. Wedemayer, J. J. Boniface, R. C. Stevens, and P. G. Schulz, *Science* 271, 1086 (1996).
160. I. Haq, *Arch. Biochem. Biophys.* 403, 1 (2002).
161. C. J. McGurk, P. J. McHugh, M. J. Tilby, K. A. Grimaldi, and J. A. Hartley, *Method Enzymol.* 340, 358 (2001).
162. C. Bailly and M. J. Waring, *Method Enzymol.* 340, 485 (2001).
163. I. Haq, T. C. Jenkins, B. Z. Chowdhry, J. S. Ren, and J. B. Chaires, *Method Enzymol.* 323, 373 (2000).
164. D. S. Pilch, N. Poklar, E. E. Baird, P. B. Dervan, and K. J. Breslauer, *Biochemistry* 38, 2143 (1999).
165. M. L. Kopka, C. Yoon, D. Goodsell, P. Pjura, and R. E. Dickerson, *Proc. Natl. Acad. Sci. USA* 82, 1376 (1985).
166. J. B. Chaires, *Biopolymers* 44, 201 (1997).
167. C. H. Lin and D. J. Patel, *J. Am. Chem. Soc.* 117, 5901 (1995).
168. C. Aleman, M. C. Vega, L. Tabernero, and J. Bella, *J. Phys. Chem.* 100, 11480 (1996).
169. K. J. McConnell and D. L. Beveridge, *J. Mol. Biol.* 304, 803 (2000).
170. R. Chattopadhyaya, S. Ikuta, K. Grzeskowiak, and R. E. Dickerson, *Nature* 334, 175 (1988).
171. H. L. Ng, M. L. Kopka, and R. E. Dickerson, *Proc. Natl. Acad. Sci. USA* 97, 2035 (2000).
172. J. Pelta, F. Livolant, and J.-L. Sikorav, *J. Biol. Chem.* 271, 5656 (1996).
173. D. Jary and J.-L. Sikorav, *Biochemistry* 38, 3223 (1999).
174. J. Liu and M. Lu, *J. Biol. Chem.* 277, 48708 (2002).
175. C. D. Snow, N. Nguyen, V. S. Pande, and M. Gruebele, *Nature* 420, 102 (2002).
176. V. Daggett and A. R. Fersht, *Trends Biochem. Sci.* 28, 18 (2003).
177. C. H. Williams, T. J. Stillman, V. V. Barynin, S. E. Sedelnikova, Y. Tang, J. Green, J. R. Guest, and P. J. Artymiuk, *Nat. Struct. Biol.* 9, 447 (2002).
178. P. Chakrabarti and J. Janin, *Proteins* 47, 334 (2002).
179. H. S. Son, B. H. Hong, C.-W. Lee, S. Yun, and K. S. Kim, *J. Am. Chem. Soc.* 123, 514 (2001).
180. J. L. Klepeis and C. A. Floudas, *J. Comput. Chem.* 23, 245 (2002).
181. J. Meller and R. Elber, *Adv. Chem. Phys.* 120, 77 (2002).
182. D. Gerber and Y. Shai, *J. Mol. Biol.* 322, 491 (2002).
183. M. Egli, *Chem. Biol.* 9, 277 (2002).
184. J. H. Wang, *Trends Biochem. Sci.* 27, 122 (2002).
185. J. D. Harper and P. T. Lansbury, *Annu. Rev. Biochem.* 66, 385 (1997).
186. A. Fink, *Fold. Des.* 3, R9 (1998).
187. L. K. Tsou, C. D. Tatko, and M. L. Waters, *J. Am. Chem. Soc.* 124, 14917 (2002).
188. X. T. Zhu, X. Zhao, W. F. Burkholder, A. Gragerov, C. M. Ogata, M. E. Gottesman, and W. A. Hendrickson, *Science* 272, 1606 (1996).
189. E. Fischer, *Ber. Dtsch. Chem. Ges.* 27, 2985 (1894).
190. D. E. Koshland, *Angew. Chem. Int. Ed. Engl.* 33, 2315 (1994).
191. J.-M. Lehn, Nobel Lecture in Chemistry, 1987.
192. C. J. Pedersen, Nobel Lecture in Chemistry, 1987.
193. D. J. Cram, Nobel Lecture in Chemistry, 1987.
194. T. Liu and H.-J. Schneider, *Angew. Chem. Int. Ed. Engl.* 41, 1368 (2002).
195. C. A. Hunter, C. M. R. Low, C. Rotger, J. G. Vinter, and C. Zonta, *Proc. Natl. Acad. Sci. USA* 99, 4873 (2002).
196. H.-J. Schneider and A. K. Mohammad-Ali, in "Comprehensive Supramolecular Chemistry" (F. Vögtle, Ed.), Vol. II, p. 69. Pergamon, Oxford, 1996.
197. "Clusters of Atoms and Molecules" (H. Haberland, Ed.). Springer-Verlag, Berlin, 1994.
198. W. Eberhardt, *Surf. Sci.* 500, 242 (2002).
199. "Nanomaterials, Synthesis, Properties, and Applications" (A. S. Edelstein and R. C. Cammarata, Eds.). Institute of Physics Publishing, Bristol, 1996.
200. W. de Heer, *Rev. Mod. Phys.* 65, 611 (1993).
201. W. Knight, K. Clemenger, W. A. de Heer, W. A. Saunders, M. Y. Chou, and M. L. Cohen, *Phys. Rev. Lett.* 52, 2141 (1984).
202. V. Bonacic-Koutecky, L. Cespiva, P. Fantucci, J. Pittner, and J. Koutecky, *J. Chem. Phys.* 100, 490 (1994).
203. I. M. L. Billas, A. Chatelain, and W. A. de Heer, *Science* 265, 1682 (1994).
204. X. Li, B. Kiran, J. Li, H.-J. Zhai, and L.-S. Wang, *Angew. Chem. Int. Ed. Engl.* 41, 4786 (2002).

205. P. Pyykkö and N. Runeberg, *Angew. Chem. Int. Ed. Engl.* 41, 2174 (2002).
206. D. A. H. Cunningham, W. Vogel, H. Kageyama, S. Tsubota, and M. Haruta, *J. Catal.* 177, 1 (1998).
207. D. T. Thompson, *Gold Bull.* (*London*), 34, 56 (2001).
208. A. W. Snow and H. Wohltjen, *Chem. Mater.* 10, 947 (1998).
209. C. J. Kiely, J. Fink, M. Brust, D. Bethell, and D. J. Shiffrin, *Nature* 396, 444 (1998).
210. B. A. Korgel and D. Fitzmaurice, *Adv. Mater.* 10, 661 (1998).
211. T. Lee, J. Liu, N.-P. Chen, R. P. Andres, D. B. Janes, and R. Reifenberger, *J. Nanopart. Res.* 2, 345 (2000).
212. M. Knupfer, *Surf. Sci. Rep.* 42, 1 (2001).
213. H. Kietzmann, R. Rochow, G. Ganteför, W. Eberhardt, K. Vietze, G. Seifert, and P. W. Fowler, *Phys. Rev. Lett.* 81, 5378 (1998).
214. R. O. Jones and G. Seifert, *Phys. Rev. Lett.* 79, 443 (1997).
215. K. S. Kim, J. M. Park, J. Kim, S. B. Suh, P. Tarakeshwar, K. H. Lee, and S. S. Park, *Phys. Rev. Lett.* 84, 2425 (2000).
216. J.-H. Cho, K. S. Kim, C. T. Chan, and Z. Zhang, *Phys. Rev. B* 113408 (2001).
217. S. P. Jarvis, T. Uchihashi, T. Ishida, and H. Tokumoto, *J. Phys. Chem. B* 104, 6091 (2000).
218. C. Yannoules, U. Landman, A. Herlert, and L. Schweikhard, *Phys. Rev. Lett.* 86, 2996 (2001).
219. "Metal Clusters" (W. Ekardt, Ed.), Wiley, New York, 1999.
220. X.-B. Wang and L.-S. Wang, *Phys. Rev. Lett.* 83, 3402 (1999).
221. M. B. More, D. Ray, and P. B. Armentrout, *J. Am. Chem. Soc.* 121, 417 (1999).
222. S.-W. Lee, S. Chang, D. Kossakovski, H. Cox, and J. L. Beauchamp, *J. Am. Chem. Soc.* 121, 10152 (1999).
223. S. W. Lee, H. Cox, W. A. Goddard, and J. L. Beauchamp, *J. Am. Chem. Soc.* 122, 9201 (2000).
224. B. S. Fox, M. K. Beyer, and V. E. Bondybey, *J. Am. Chem. Soc.* 124, 13613 (2002).
225. H. M. Lee, J. Kim, S. Lee, B. J. Mhin, and K. S. Kim, *J. Chem. Phys.* 111, 3995 (1999).
226. H. M. Lee, S. B. Suh, J. Y. Lee, P. Tarakeshwar, and K. S. Kim, *J. Chem. Phys.* 112, 9759 (2000).
227. H. M. Lee, S. B. Suh, and K. S. Kim, *J. Chem. Phys.* 114, 10749 (2001).
228. K. S. Kim, M. Dupuis, G. C. Lie, and E. Clementi, *Chem. Phys. Lett.* 131, 451 (1986).
229. P. Tarakeshwar, K. S. Kim, S. Djafari, K. Buchhold, B. Reimann, H.-D. Barth, and B. Brutschy, *J. Chem. Phys.* 114, 4016 (2001).
230. S. Park, D. Srivastava, and K. Cho, *Nanotechnology* 12, 245 (2001).
231. I. Park, K. Cho, S. Lee, K. S. Kim, and J. D. Joannopoulos, *Comput. Mater. Sci.* 21, 291 (2001).
232. K. S. Kim, P. Tarakeshwar, and J. Y. Lee, *Chem. Rev.* 100, 4145 (2000).
233. H. G. Kim, C.-W. Lee, S. Yun, B. H. Hong, Y.-O. Kim, D. Kim, H. Ihm, J. W. Lee, E. C. Lee, P. Tarakeshwar, S.-M. Park, and K. S. Kim, *Org. Lett.* 4, 3971 (2002).
234. C. J. Pedersen, *J. Am. Chem. Soc.* 89, 2495 (1967).
235. C. J. Pedersen, *J. Am. Chem. Soc.* 89, 7017 (1967).
236. B. Dietrich, J.-M. Lehn, and J.-P. Sauvage, *Tetrahedron Lett.* 10, 2885 (1969).
237. B. Dietrich, J.-M. Lehn, and J.-P. Sauvage, *Tetrahedron Lett.* 10, 2885 (1969).
238. G. W. Gokel and D. J. Cram, *J. Chem. Soc. Chem. Commun.* 481, (1973).
239. D. J. Cram and J. M. Cram, *Science* 183, 803 (1974).
240. "Comprehensive Supramolecular Chemistry" (G. W. Gokel, Ed.), Vol. I. Pergamon, Oxford, 1996.
241. "Comprehensive Supramolecular Chemistry" (F. Vögtle, Ed.), Vol. II. Pergamon, Oxford, 1996.
242. G. W. Gokel and O. Murillo, in "Comprehensive Supramolecular Chemistry" (G. W. Gokel, Ed.), Vol. I. p. 1. Pergamon, Oxford, 1996.
243. J. S. Bradshaw, R. M. Izatt, A. V. Bordunov, C. Y. Zhu, and J. K. Hathway, in "Comprehensive Supramolecular Chemistry" (G. W. Gokel, Ed.), Vol. I, p. 35. Pergamon, Oxford, 1996.
244. R. M. Izatt, K. Pawlak, and J. S. Bradshaw, *Chem. Rev.* 91, 1721 (1991).
245. R. M. Izatt, K. Pawlak, and J. S. Bradshaw, *Chem. Rev.* 95, 2529 (1995).
246. X. X. Zhang, R. M. Izzat, J. S. Bradshaw, and K. E. Krakowiak, *Coord. Chem. Rev.* 174, 179 (1998).
247. J. W. Steed, *Coord. Chem. Rev.* 215, 171 (2001).
248. B. Dietrich, in "Comprehensive Supramolecular Chemistry" (G. W. Gokel, Ed.), Vol. I, p. 153. Pergamon, Oxford, 1996.
249. E. Maverick and D. J. Cram, in "Comprehensive Supramolecular Chemistry" (G. W. Gokel, Ed.), Vol. I, p. 213. Pergamon, Oxford, 1996.
250. M. A. McKervey, M.-J. Schwing-Weill, and F. Arnaud-Neu, in "Comprehensive Supramolecular Chemistry" (G. W. Gokel, Ed.), Vol. I, p. 537. Pergamon, Oxford, 1996.
251. J. Telford and K. N. Raymond, in "Comprehensive Supramolecular Chemistry" (G. W. Gokel, Ed.), Vol. I, p. 245. Pergamon, Oxford, 1996.
252. M. Dobler, in "Comprehensive Supramolecular Chemistry" (G. W. Gokel, Ed.), Vol. I, p. 267. Pergamon, Oxford, 1996.
253. "Comprehensive Supramolecular Chemistry" (D. N. Reinhoudt, Ed.), Vol. X, Pergamon, Oxford, 1996.
254. D.-J. van Unen, J. F. J. Engbersen, and D. N. Reinhoudt, *Biotechnol. Bioeng.* 77, 248 (2002).
255. J. S. Bradshaw and R. M. Izatt, *Acc. Chem. Res.* 30, 338 (1997).
256. H. S. Choi, D. Kim, P. Tarakeshwar, S. B. Suh, and K. S. Kim, *J. Org. Chem.* 67, 1848 (2002).
257. A. Pochini and R. Ungaro, in "Comprehensive Supramolecular Chemistry" (F. Vögtle, Ed.), Vol. II, p. 103. Pergamon, Oxford, 1996.
258. D. A. Dougherty, in "Comprehensive Supramolecular Chemistry" (F. Vögtle, Ed.), Vol. II, p. 195. Pergamon, Oxford, 1996.
259. F. Vögtle, C. Seel, and P.-M. Windscheif, in "Comprehensive Supramolecular Chemistry" (F. Vögtle, Ed.), Vol. II, Pergamon, Oxford, 1996.
260. "Comprehensive Supramolecular Chemistry" (J. Szejtli and T. Osa, Vol. III. Pergamon, Oxford, 1996.
261. K. A. Connors, *Chem. Rev.* 97, 1325 (1997).
262. W. L. Mock, in "Comprehensive Supramolecular Chemistry" (F. Vögtle, Ed.), Vol. II, p. 477. Pergamon, Oxford, 1996.
263. E. Maverick and D. J. Cram, in "Comprehensive Supramolecular Chemistry" (F. Vögtle, Ed.), Vol. II, p. 367. Pergamon, Oxford, 1996.
264. R. Dutzler, E. B. Campbell, M. Cadene, B. T. Chait, and R. Mackinnon, *Nature* 415, 287 (2002).
265. D. A. Doyle, J. M. Cabral, R. A. Pfuetzner, A. Kuo, J. M. Gulbis, S. L. Cohen, B. T. Chait, and R. MacKinnon, *Science* 280, 69 (1998).
266. B. Roux and R. Mackinnon, *Science* 285, 100 (1999).
267. J. Åqvist and V. Luzhkov, *Nature* 404, 881 (2000).
268. W. Verboom and D. N. Reinhoudt, in "Comprehensive Supramolecular Chemistry" (F. Vögtle, Ed.), Vol. II, p. 495. Pergamon, Oxford, 1996.
269. M. T. Reetz, in "Comprehensive Supramolecular Chemistry" (F. Vögtle, Ed.), Vol. II, p. 563. Pergamon, Oxford, 1996.
270. K. Müller-Dethlefs and P. Hobza, *Chem. Rev.* 100, 143 (2000).
271. B. Jeziorski, R. Moszynski, and K. Szalewicz, *Chem. Rev.* 94, 1887 (1994).
272. J. N. Israelachvili, "Intermolecular and Surface Forces." Academic Press, London, 1992.
273. C. C. J. Roothan, *Rev. Mod. Phys.* 23, 69 (1951).
274. C. Møller and M.-S. Plesset, *Phys. Rev.* 46, 618 (1934).
275. D. Cremer, in "Encyclopedia of Computational Chemistry" (P. v. R. Schleyer, Ed.), p. 1706. Wiley, New York, 1998.

276. J. C. H. Spence, *Mater. Sci. Eng. R* 26, 1 (1999).
277. S. Iijima, *Nature* 354 56 (1991).
278. D. J. Muller, H. Janovjak, T. Lehto, L. Kuerschner, and K. Anderson, *Progr. Biophys. Mol. Biol.* 79, 1 (2002).
279. R. Garcia and R. Perez, *Surf. Sci. Rep.* 47, 197 (2002).
280. P. A. Rock, "Chemical Thermodynamics." University Science Books, Herndon, 1983.
281. H. Senderowitz, D. Q. McDonald, and W. C. Still, *J. Org. Chem.* 62, 9123 (1997).
282. K. Szalewicz and B. Jeziorski, in "Molecular Interactions—From van der Waals to Strongly Bound Complexes" (S. Scheiner, Ed.), p. 3. Wiley, New York, 1997.
283. B. Jeziorski, R. Moszynski, A. Ratkiewicz, S. Rybak, K. Szalewicz, and H. L. Williams, in "Methods and Techniques in Computational Chemistry: METECC-94, (E. Clementi Ed.), Vol. B, Medium Size Systems, p. 79. STEF, Cagliari, 1993.
284. G. C. Pimentel and A. C. McClellan, "The Hydrogen Bond." Freeman, San Francicso, 1960.
285. J. A. Jeffrey and W. Saenger, "Hydrogen Bonding in Biological Structures." Springer-Verlag, Berlin, 1991.
286. J. A. Jeffrey, "Introduction to Hydrogen Bonding." University Press, Oxford, 1997.
287. S. Scheiner, "Hydrogen Bonding: A Theoretical Perspective." University Press, Oxford, 1997.
288. J. D. Dunitz and R. Taylor, *Chem. Eur. J.* 3, 89 (1997).
289. L. F. Scatena, M. G. Brown, and G. L. Richmond, *Science* 292, 908 (2001).
290. D. Chamma and O. Henri-Rousseau, *Chem. Phys.* 248, 53 (1999).
291. P. Auffinger and E. Westhof, *J. Mol. Biol.* 300, 1113 (2000).
292. B. Schneider, K. Patel, and H. M. Berman, *Biophys. J.* 75, 2422 (1998).
293. E. T. Kool, *Chem. Rev.* 97, 1473 (1997).
294. C. L. Perrin and J. B. Nielson, *Annu. Rev. Phys. Chem.* 48, 511 (1997).
295. T. S. Moore and T. F. Winmill, *J. Chem. Soc.* 101, 1635 (1912).
296. W. M. Latimer and W. H. Rodebish, *J. Am. Chem. Soc.* 42, 1419 (1920).
297. L. Pauling, *J. Am. Chem. Soc.* 53, 1367 (1931).
298. L. Pauling, in "Hydrogen Bonding" (D. Hadzi and H. W. Thompson, Eds.), p. 1. Pergamon Press, New York, 1959.
299. T. A. Andrea, W. C. Swope, and H. C. Andersen, *J. Chem. Phys.* 79, 4576 (1984).
300. C. H. Cho, S. Singh, and G. W. Robinson, *Faraday Discuss.* 103, 19 (1996).
301. P. G. Kusalik and I. M. Svishchev, *Science* 265, 1219 (1994).
302. D. Eisenberg and W. Kauzman, "The Structure and Properties of Water." Oxford Univ. Press, London, 1969.
303. L. J. Barbour, G. W. Orr, and J. L. Atwood, *Nature* 393, 671 (1998).
304. J. Urquidi, S. Singh, C. H. Cho, and G. W. Robinson, *Phys. Rev. Lett.* 83, 2348 (1999).
305. K. Hermansson, *Chem. Phys. Lett.* 260, 229 (1996).
306. P. Jedlovszky, M. Mezei, and R Vallauri, *Chem. Phys. Lett.* 318, 155 (2000).
307. D. Peeters, *J. Mol. Liquids* 67, 49 (1995).
308. P. L. Silvestrelli and M. Parrinello, *J. Chem. Phys.* 111, 3572 (1999).
309. D. J. Wales and M. P. Hodges, *Chem. Phys. Lett.* 286, 65 (1998).
310. K. S. Kim, B. J. Mhin, U.-S. Choi, and K. Lee, *J. Chem. Phys.* 97, 6649 (1992).
311. J. Kim, J. Y. Lee, S. Lee, B. J. Mhin, and K. S. Kim, *J. Chem. Phys.* 102, 310 (1995).
312. B. J. Mhin, S. J. Lee, and K. S. Kim, *Phys. Rev. A* 48, 3764 (1993).
313. B. J. Mhin, J. Kim, S. Lee, J. Y. Lee, and K. S. Kim, *J. Chem. Phys.* 100, 4484 (1994).
314. J. Kim, B. J. Mhin, S. J. Lee, and K. S. Kim, *Chem. Phys. Lett.* 219, 243 (1994).
315. J. Kim and K. S. Kim, *J. Chem. Phys.* 109, 5886 (1998).
316. J. Kim, D. Majumdar, H. M. Lee, and K. S. Kim, *J. Chem. Phys.* 110, 9218 (1999).
317. J. P. Devlin, C. Joyce, and V. Buch, *J. Phys. Chem. A* 104, 1974 (2000).
318. R. Ludwig, *Angew. Chem. Int. Ed. Engl.* 40, 1808 (2001).
319. T. Head-Gordon and G. Hura, *Chem. Rev.* 102, 2651 (2002).
320. J.-H. Guo, Y. Luo, A. Augustsson, J.-E. Rubensson, C. Såthe, C. Ågren, H. Siegbahn, and J. Nordgren, *Phys. Rev. Lett.* 89, 137402 (2002).
321. T. K. Ghanty, V. N. Staroverov, P. R. Koren, and E. R. Davidson, *J. Am. Chem. Soc.* 122, 1210 (2000).
322. Y. Ikezoe, N. Hirota, J. Nakagawa, and K. Kitazawa, *Nature* 393, 749 (1998).
323. M. F. Kropman and H. J. Bakker, *Science* 291, 2118 (2001).
324. A. Milet, R. Moszynski, P. E. S. Wormer, and A. van der Avoird, *J. Phys. Chem. A* 103, 6811 (1999).
325. G. C. Groenenboom, E. M. Mas, R. Bukowski, K. Szalewicz, P. E. S. Wormer, and A. van der Avoird, *Phys. Rev. Lett.* 84, 4072 (2000).
326. Y. Gu, T. Kar, and S. Scheiner, *J. Am. Chem. Soc.* 121, 9411 (1999).
327. B. Reimann, K. Buchhold, S. Vaupel, B. Brutschy, Z. Havlas, V. Spirko, and P. Hobza, *J. Phys. Chem. A* 105, 5560 (2001).
328. P. Hobza, V. Spirko, Z. Havlas, K. Buchhold, B. Reimann, H.-D. Barth, and B. Brutschy, *Chem. Phys. Lett.* 299, 180 (1999).
329. M. Brandl, M. Meyer, and J. Sühnel, *J. Biomol. Struct. Dyn.* 18, 545 (2001).
330. J. L. Atwood, L. J. Barbour, A. Jerga, and B. L. Schottel, *Science* 298, 1000 (2002).
331. M. C. Etter, J. MacDonald, and J. Bernstein, *Acta Cryst. B* 46, 256 (1991).
332. C. S. Wilcox, E. Kim, D. Romano, L. H. Kuo, A. L. Burt, and D. P. Curran, *Tetrahedron* 51, 621 (1995).
333. R. Deans, G. Cooke, and V. M. Rotello, *J. Org. Chem.* 62, 836 (1997).
334. G. Cooke and V. M. Rotello, *Chem. Soc. Rev.* 31, 275 (2002).
335. S. U. Chang and A. D. Hamilton, *J. Am. Chem. Soc.* 110, 1318 (1988).
336. F. H. Beijer, R. P. Sijbesma, J. A. J. M. Vekemans, E. W. Meijer, H. Kooijman, and A. L. Spek, *J. Org. Chem.* 61, 6371 (1996).
337. A. R. A. Palmans, J. A. J. M. Vekemans, H. Fischer, R. A. Hikmet, and E. W. Meijer, *Chem. Eur. J.* 3, 300 (1997).
338. R. F. M. Lange, F. H. Beijer, R. P. Sijbesma, R. W. W. Hooft, H. Kooijman, A. L. Spek, J. Kroon, and E. W. Meijer, *Angew. Chem. Int. Ed. Engl.* 36, 969 (1997).
339. R. P. Sijbesma, F. H. Beijer, L. Brunsveld, B. J. B. Folmer, J. Hirschberg, R. F. M. Lange, J. K. L. Lowe, and E. W. Meijer, *Science* 278, 1601 (1997).
340. F. H. Beijer, H. Kooijman, A. L. Spek, R. P. Sijbesma, and E. W. Meijer, *Angew. Chem. Int. Ed. Engl.* 37, 75 (1998).
341. B. J. B. Folmer, R. P. Sijbesma, H. Kooijman, A. L. Spek, and E. W. Meijer, *J. Am. Chem. Soc.* 121, 9001 (1999).
342. R. F. M. Lange, M. v. Gurp, and E. W. Meijer, *J. Poly. Sci. A.* 3657 (1999).
343. S. H. M. Söntjens, R. P. Sijbesma, M. H. P. v. Genderen, and E. W. Meijer, *J. Am. Chem. Soc.* 122, 7487 (2000).
344. J. H. K. K. Hirschberg, L. Brunsveld, A. Ramzi, J. A. J. M. Vekemans, R. P. Sijbesma, and E. W. Meijer, *Nature* 407, 167 (2000).
345. A. El-ghayoury, E. Peeters, A. P. H. J. Schenning, and E. W. Meijer, *J. Chem. Soc. Chem. Commun.* 1969 (2000).
346. J. Sartorius and H.-J. Schneider, *Chem. Eur. J.* 2, 1446 (1996). Quadruple bonds are more stable than triple.
347. P. S. Corbin and S. C. Zimmerman, *J. Am. Chem. Soc.* 120, 9710 (1998).
348. M. S. Lüth, E. Freisinger, and B. Lippert, *Chem. Eur. J.* 7, 2104 (2001).
349. M. P. Hughes and B. D. Smith, *J. Org. Chem.* 62, 4492 (1997).

350. J. H. R. Tucker and S. Collinson, *Chem. Soc. Rev.* 31, 147 (2002).
351. Y. Ge, L. Miller, T. Ouimet, and D. K. Smith, *J. Org. Chem.* 65, 8831 (2000).
352. A. O. Cuello, C. M. McIntosh, and V. M. Rotello, *J. Am. Chem. Soc.* 122, 3517 (2000).
353. A. Niemz and V. M. Rotello, *Acc. Chem. Res.* 32, 44 (1999).
354. A. K. Boal and V. M. Rotello, *J. Am. Chem. Soc.* 121, 4914 (1999).
355. E. Breinlinger, A. Niemz, and V. M. Rotello, *J. Am. Chem. Soc.* 117, 5379 (1995).
356. A. M. Brouwer, C. Frochot, F. G. Gatti, D. A. Leigh, L. Mottier, F. Paolucci, S. Roffia, and G. W. H. Wurpel, *Science* 291, 2124 (2001).
357. A. Goodman, E. Breinlinger, M. Ober, and V. M. Rotello, *J. Am. Chem. Soc.* 123, 6213 (2001).
358. S. Goswami, K. Ghosh, and M. Halder, *Tetrahedron Lett.* 40, 1735 (1999).
359. R. D. Shannon, *Acta Crystallogr. A* 32, 751 (1976).
360. D. Marx, M. E. Tuckerman, J. Hutter, and M. Parrinello, *Nature* 397, 601 (1999).
361. E. A. Steel, K. M. Merz, A. Selinger, and A. W. Castleman, *J. Phys. Chem.* 99, 7829 (1995).
362. M. P. Hodges and D. J. Wales, *Chem. Phys. Lett.* 324, 279 (2000).
363. B. J. Mhin, J. Kim, and K. S. Kim, *Chem. Phys. Lett.* 216, 305 (1993).
364. K. S. Kim, S. Lee, B. J. Mhin, S. J. Cho, and K. S. Kim, *Chem. Phys. Lett.* 216, 309 (1993).
365. S. Lee, J. Kim, J. K. Park, and K. S. Kim, *J. Phys. Chem.* 100, 14329 (1996).
366. J. Baik, J. Kim, D. Majumdar, and K. S. Kim, *J. Chem. Phys.* 110, 9116 (1999).
367. J. Kim, H. M. Lee, S. B. Suh, D. Majumdar, and K. S. Kim, *J. Chem. Phys.* 113, 5259 (2000).
368. H. M. Lee, D. Kim, and K. S. Kim, *J. Chem. Phys.* 116, 5509 (2002).
369. H. M. Lee and K. S. Kim, *J. Chem. Phys.* 114, 4461 (2001).
370. H. J. Bakker and H.-K. Nienhuys, *Science* 297, 587 (2002).
371. M. Kaupp and P. v. R. Schleyer, *J. Am. Chem. Soc.* 96, 7316 (1992).
372. D. Feller, E. D. Glendening, R. A. Kendall, and K. A. Peterson, *J. Chem. Phys.* 100, 4981 (1994).
373. A. L. Derepas, J.-M. Soudan, V. Brenner, J.-P. Dognon, and Ph. Millié, *J. Comput. Chem.* 23, 1013 (2002).
374. H. L. Loeffler and B. M. Rode, *J. Chem. Phys.* 117, 110 (2002).
375. E. M. Cabaleiro-Lago and J. Rodríguez-Otero, *J. Phys. Chem. A* 106, 7195 (2002).
376. T. D. Vaden, B. Forinash, and J. Lisy, *J. Chem. Phys.* 117, 4628 (2002).
377. M. Okumura, L. I. Yeh, J. D. Myers, and Y. T. Lee, *J. Phys. Chem.* 94, 3416 (1990).
378. K. S. Oh, C.-W. Lee, H. S. Choi, S. J. Lee, and K. S. Kim, *Org. Lett.* 2, 2679 (2000).
379. S. Yun, Y.-O. Kim, D. Kim, H. Kim, H. Ihm, J. K. Kim, C.-W. Lee, W. J. Lee, J. Yoon, K. S. Oh, J. Yoon, S.-M. Park, and K. S. Kim, *Org. Lett.* 5, 471 (2003).
380. C. Cui, S. J. Cho, and K. S. Kim, *J. Phys. Chem. A* 102, 1119 (1998).
381. F. F. Schmidtchen and M. Berger, *Chem. Rev.* 97, 1609 (1997).
382. R. L. E. Furlan, Y.-F. Ng, G. R. L. Cousins, J. R. Redman, and J. K. M. Sanders, *Tetrahedron* 58, 771 (2002).
383. H. Ihm, S. Yun, H. G. Kim, J. K. Kim, and K. S. Kim, *Org. Lett.* 4, 2897 (2002).
384. A. Marquis, J.-P. Kintzinger, R. Graff, P. N. W. Baxter, and J.-M. Lehn, *Angew. Chem. Int. Ed. Engl.* 41, 2760 (2002).
385. T. Nabeshima, *Coord. Chem. Rev.* 148, 151 (1996).
386. G. Das, H. Onouchi, E. Yashima, N. Sakai, and S. Matile, *Chem. Biol. Chem.* 3, 1089 (2002).
387. R. Fiammengo, P. Timmerman, J. Huskens, K. Versluis, A. J. R. Heck, and D. N. Reinhoudt, *Tetrahedron* 58, 757 (2002).
388. T. B. Norsten, K. Chichak, and N. R. Branda, *Tetrahedron* 58, 639 (2002).
389. M. Cai, X. Shi, V. Sidorov, D. Fabris, Y.-F. Lam, and J. T. Davis, *Tetrahedron* 58, 661 (2002).
390. R. Nonokawa and E. Yashima, *J. Am. Chem. Soc.* 125, 1278 (2003).
391. A. M. Klibanov, *Trends Biotechnol.* 15, 97 (1997).
392. "Enzymatic Reactions in Organic Media" (A. M. Klibanov and A. M. P. Koskinen, Eds.). Chapman & Hall, Glassgow, 1996.
393. H. Tsukube, A. Betchaku, Y. Hiyama, and T. Itoh, *J. Org. Chem.* 59, 7014 (1994).
394. J. Broos, A. J. W. G. Visser, J. F. J. Engbersen, W. Verboom, and D. N. Reinhoudt, *J. Am. Chem. Soc.* 117, 12657 (1995).
395. J. Broos, I. K. Sakodinskaya, J. F. J. Engbersen, W. Verboom, and D. N. Reinhoudt, *J. Chem. Soc., Chem. Commun.* 255 (1995).
396. J. Broos, R. Arend, G. B. van Dijk, W. Verboom, J. F. J. Engbersen, and D. N. Reinhoudt, *J. Chem. Soc., Perkin Trans.* 1415 (1996).
397. J. F. J. Engbersen, J. Broos, W. Verboom, and D. N. Reinhoudt, *Pure Appl. Chem.* 68, 2171 (1996).
398. D.-J. van Unen, J. F. J. Engbersen, and D. N. Reinhoudt, *Biotechnol. Bioeng.* 59, 553 (1998).
399. D.-J. van Unen, I. K. Sakodinskaya, J. F. J. Engbersen, and D. N. Reinhoudt, *J. Chem. Soc., Perkin Trans.* 3343 (1998).
400. D.-J. van Unen, J. F. J. Engbersen, and D. N. Reinhoudt, *J. Mol. Catal. B* 11, 877 (2000).
401. W. W. Cleland, *Biochemistry* 31, 317 (1992).
402. W. W. Cleland and M. M. Krevoy, *Science* 264, 1887 (1994).
403. J. A. Gerlt and P. G. Grassman, *Biochemistry* 32, 11943 (1993).
404. J. A. Gerlt and P. G. Grassman, *J. Am. Chem. Soc.* 115, 11552 (1993).
405. P. A. Frey, S. A. Whitt, and J. B. Tobin, *Science* 264, 1927 (1994).
406. Y.-J. Zheng and T. C. Bruice, *Proc. Natl. Acad. Sci. USA* 94, 4285 (1997).
407. P. A. Frey and W. W. Cleland, *Biorg. Chem.* 26, 175 (1998).
408. B. Schiøtt, B. B. Iversen, G. K. H. Madsen, F. K. Larsen, and T. C. Bruice, *Proc. Natl. Acad. Sci. USA* 95, 12799 (1998).
409. A. Warshel, A. Papazyan, and P. A. Kollman, *Science* 269, 102 (1995).
410. W. W. Cleland and M. M. Krevoy, *Science* 269, 104 (1995).
411. P. A. Frey, *Science* 269, 104 (1995).
412. N. C. Ha, G. Choi, K. Y. Choi, and B. H. Oh, *Curr. Opin. Struct. Biol.* 11, 674 (2001).
413. H. S. Cho, N. C. Ha, G. Choi, H. J. Kim, D. Lee, K. S. Oh, K. S. Kim, W. Lee, K. Y. Choi, and B. H. Oh, *J. Biol. Chem.* 274, 32863 (1999).
414. V. A. Russell, M. C. Etter, and M. D. Ward, *J. Am. Chem. Soc.* 116, 1941 (1994).
415. V. A. Russell, C. C. Evans, W. Li, and M. D. Ward, *Science* 276, 575 (1997).
416. V. A. Russell and M. D. Ward, *J. Mater. Chem.* 7, 1123 (1997).
417. V. A. Russell and M. D. Ward, *New J. Chem.* 149 (1998).
418. X. Bu, P. Feng, and G. D. Stucky, *Science* 278, 2080 (1997).
419. L. B. McCusker, Ch. Baerlocher, R. Grosse-Kunstleve, S. Brenner, and T. Wessels, *Chimia* 55, 497 (2001).
420. K. R. Seddon and C. B. Aäkeroy, *Chem. Soc. Rev.* 397 (1993).
421. J.-M. Lehn, A. Rigault, J. Siegel, J. Harrowfield, B. Chevrier, and D. Moras, *Proc. Natl. Acad. Sci. USA* 27, 2565 (1987).
422. A. Marquis-Rigault, A. Dupont-Gervais, A. van Dorssealer, and J.-M. Lehn, *Chem. Eur. J.* 3, 1395 (1996).
423. T. M. Garrett, U. Koert, J.-M. Lehn, A. Rigault, D. Meyer, and J. Fischer, *J. Chem. Soc., Chem. Commun.* 557 (1990).
424. R. Kramer, J.-M. Lehn, and A. Marquis-Rigault, *Proc. Natl. Acad. Sci. USA* 90, 5394 (1993).
425. Y. Inoue, T. Hakushi, Y. Liu, L.-H. Tong, B.-J. Shen, and D.-S. Jin, *J. Am. Chem. Soc.* 115, 475 (1993).
426. D. W. Pratt, *Annu. Rev. Phys. Chem.* 49, 481 (1998).
427. P. Koehl and M. Levitt, *J. Mol. Biol.* 293, 1161 (1999).
428. P. Mombaerts, *Science* 286, 707 (1999).
429. T. Ren, Y. Jin, K. S. Kim, and D. H. Kim, *J. Biomol. Struct. Dyn.* 15, 401 (1997).

430. B. H. Hong, J. Y. Lee, S. J. Cho, S. Yun, and K. S. Kim, *J. Org. Chem.* 64, 5661 (1999).
431. F. H. Allen, J. A. K. Howard, V. J. Hoy, G. R. Desiraju, D. S. Reddy, and C. C. Wilson, *J. Am. Chem. Soc.* 118, 4081 (1996).
432. J. C. Ma and D. A. Dougherty, *Chem. Rev.* 97, 1303 (1997).
433. P. Privalov and S. J. Gill, *Adv. Protein Chem.* 40, 191 (1988).
434. S. L. Burley and G. A. Petsko, *Science* 229, 23 (1985).
435. A. Ikeda, C. Fukuhara, and S. Shinkai, *Tetrahedron Lett.* 37, 7091 (1996).
436. O. P. Kryatova, A. G. Kolchinski, and E. V. Rybak-Akimova, *Tetrahedron* 59, 231 (2003).
437. C. A. Hunter and J. K. M. Sanders, *J. Am. Chem. Soc.* 112, 5525 (1990).
438. K. S. Kim, J. Y. Lee, S. J. Lee, T.-K. Ha, and D. Kim, *J. Am. Chem. Soc.* 116, 7399 (1994).
439. J. Y. Lee, S. J. Lee, H. S. Choi, S. J. Cho, K. S. Kim, and T.-K. Ha, *Chem. Phys. Lett.* 232, 67 (1995).
440. K. S. Kim, J. Y. Lee, H. S. Choi, J. Kim, and J. H. Jang, *Chem. Phys. Lett.* 265, 497 (1997).
441. P. Tarakeshwar, S. J. Lee, J. Y. Lee, and K. S. Kim, *J. Chem. Phys.* 108, 7217 (1998).
442. P. Tarakeshwar, S. J. Lee, J. Y. Lee, and K. S. Kim, *J. Phys. Chem. B* 103, 184 (1999).
443. P. Tarakeshwar, K. S. Kim, and B. Brutschy, *J. Chem. Phys.* 110, 8501 (1999).
444. P. Tarakeshwar, H. S. Choi, S. J. Lee, J. Y. Lee, K. S. Kim, T.-K. Ha, J. H. Jang, J. G. Lee, and H. Lee, *J. Chem. Phys.* 111, 5838 (1999).
445. P. Tarakeshwar, K. S. Kim, and B. Brutschy, *J. Chem. Phys.* 112, 1769 (2000).
446. P. Tarakeshwar, K. S. Kim, and B. Brutschy, *J. Chem. Phys.* 114, 1295 (2001).
447. P. Tarakeshwar, H. S. Choi, and K. S. Kim, *J. Am. Chem. Soc.* 123, 3323 (2001).
448. P. Tarakeshwar, K. S. Kim, E. Kraka, and D. Cremer, *J. Chem. Phys.* 115, 6018 (2001).
449. J. M. Park, P. Tarakeshwar, K. S. Kim, and T. Clark, *J. Chem. Phys.* 116, 10684 (2002).
450. D. Kim, S. Hu, P. Tarakeshwar, K. S. Kim, and J. M. Lisy, *J. Phys. Chem. A* 107, 1228 (2003).
451. M. O. Sinnokrot, E. F. Valeev, and C. D. Sherrill, *J. Am. Chem. Soc.* 124, 10887 (2002).
452. H. J. Neusser and H. Krause, *Chem. Rev.* 94, 1829 (1994).
453. P. Hobza, H. L. Selzle, and E. Schlag, *Chem. Rev.* 94, 1767 (1994).
454. J. C. Amicangelo and P. B. Armentrout, *J. Phys. Chem. A* 104, 11420 (2000).
455. J. E. Del Bene and I. Cohen, *J. Am. Chem. Soc.* 100, 5285 (1978).
456. K. B. Wiberg, C. M. Hadad, T. J. LePage, C. M. Breneman, and M. J. Frisch, *J. Phys. Chem.* 96, 671 (1992).
457. D. Quinonero, C. Garau, C. Rotger, A. Frontera, P. Ballester, A. Costa, and P. M. Deya, *Angew. Chem. Int. Ed. Engl.* 41, 3389 (2002).
458. I. Alkorta, I. Rozas, and J. Elguero, *J. Am. Chem. Soc.* 124, 8593 (2002).
459. M. Mascal, A. Armstrong, and M. D. Bartberger, *J. Am. Chem. Soc.* 124, 6274 (2002).
460. K. Le Barbu, J. Schiedt, R. Weinkauf, E. W. Schlag, J. M. Nilles, S. J. Xu, O. C. Thomas, and K. H. Bowen, *J. Chem. Phys.* 116, 9663 (2002).
461. A. F. Jalbout and L. Adamowicz, *J. Chem. Phys.* 116, 9672 (2002).
462. T. Maeyama, T. Oikawa, K. Seguchi, and N. Mikami, *J. Phys. Chem. A* 101, 8371 (1997).
463. J. K. Kochi, *Adv. Phys. Org. Chem.* 29, 185 (1994).
464. G. A. Olah, *Angew. Chem. Int. Ed. Engl.* 34, 1393 (1995).
465. P. Tarakeshwar, J. Y. Lee, and K. S. Kim, *J. Phys. Chem. A* 102, 2253 (1998).
466. P. Tarakeshwar and K. S. Kim, *J. Phys. Chem. A* 103, 9116 (1999).
467. C. Janiak, *J. Chem. Soc., Dalton Trans.* 3885 (2000).
468. A. P. West, S. Mecozzi, and D. A. Dougherty, *J. Phys. Org. Chem.* 10, 347 (1997).
469. M. I. Cabaço, Y. Danten, M. Besnard, Y. Guissani, and B. Guillot, *J. Phys. Chem. B* 102, 10712 (1998).
470. D. Philip and J. M. A. Robinson, *J. Chem. Soc. Perkin Trans.* 2, 1643 (1998).
471. G. Maier and Ch. Lautz, *Eur. J. Org. Chem.* 9, 769 (1998).
472. S. M. Resende and W. B. De Almeida, *Chem. Phys.* 206, 1 (1996).
473. A. Karpfen, *J. Phys. Chem. A* 103, 11431 (1999).
474. S. Tsuzuki, T. Uchimaru, and K. Tanabe, *Chem. Phys. Lett.* 287, 202 (1998).
475. M. C. Chan, P. A. Block, and R. E. Miller, *J. Chem. Phys.* 102, 3993 (1995).
476. P. Hobza, O. Bludsky, H. L. Selzle, and E. W. Schlag, *J. Chem. Phys.* 98, 6223 (1993).
477. J. T. Su and A. H. Zewail, *J. Phys. Chem. A* 102, 4082 (1998).
478. J. M. Cram and J. M. Collet, "Container Molecules and Their Guests." Royal Society of Chemistry, Cambridge, UK, 1994.
479. U. Lücking, F. C. Tucci, D. M. Rudkevich, and J. M. Rebek, *J. Am. Chem. Soc.* 122, 8880 (2000).
480. C. L. D. Gibb, H. Xi, P. A. Politzer, M. Concha, and B. C. Gibb, *Tetrahedron* 58, 673 (2002).
481. D. Chakarov and B. Kasemo, *Phys. Rev. Lett.* 81, 5181 (1998).
482. M. Pérez-Alvarez, F. M. Raymo, S. J. Rowan, D. Schiraldi, J. F. Stoddart, Z.-H. Wang, A. J. P. White, and D. J. Williams, *Tetrahedron* 57, 3799 (2001).
483. C. A. Hunter and R. Tregonning, *Tetrahedron* 58, 691 (2002).
484. H.-J. Schneider, L. Tianjin, M. Sirish, and V. Malinovski, *Tetrahedron* 58, 779 (2002).
485. M. C. T. Fyfe, J. F. Stoddart, A. J. P. White, and D. J. Williams, *New J. Chem.* 155 (1998).
486. R. P. Sijbesma, S. S. Wijmenga, and R. J. M. Nolte, *J. Am. Chem. Soc.* 114, 9807 (1992).
487. P. Pyykkö, *Chem. Rev.* 97, 597 (1997).
488. R. Landauer, *J. Phys. Condens. Matter* 1, 8099 (1989).
489. N. Agraït, A. L. Yeyati, and J. M. van Ruitenbeek, *Phys. Rep*. 377, 81 (2003).
490. A. Codina, E. J. Fernández, P. G. Jones, A. Laguna, J. M. López-de-Luzuriaga, M. Monge, M. E. Olmos, J. Pérez, and M. A. Rodriguez, *J. Am. Chem. Soc.* 124, 6781 (2002).
491. H. Schmidbaur, A. Hamel, N. W. Mitzel, A. Schier, and S. Nogai, Proc. *Natl. Acad. Sci. USA* 99, 4916 (2002).
492. H. Schmidbaur, *Nature* 413, 31 (2001).
493. H. Schmidbaur, *Gold Bull.* 33, 3 (2000).
494. R. E. Bachman, M. S. Fioritto, S. K. Fetics, and T. M. Cocker, *J. Am. Chem. Soc.* 123, 5376 (2001).
495. P. Pyykkö, W. Schneider, A. Bauer, A. Bayler, and H. Schmidbaur, *J. Chem. Soc., Chem. Commun.* 1111 (1997).
496. T. Mathieson, A. Schier, and H. Schmidbaur, *J. Chem. Soc. Dalton Trans.* 3881 (2000).
497. H. Schmidbaur, *Chem. Soc. Rev.* 24, 391 (1995).
498. M. Bardaji, P. Uznanski, C. Amiens, B. Chaudret, and A. Laguna, *J. Chem. Soc., Chem. Commun.* 598 (2002).
499. M. A. Rawashdeh-Omary, M. A. Omary, and J. P. Fackler, *Inorg. Chim. Acta* 334, 376 (2002).
500. M. A. Rawashdeh-Omary, M. A. Omary, and H. H. Patterson, *J. Am. Chem. Soc.* 122, 10371 (2002).
501. D. B. Leznoff, B.-Y. Xue, R. J. Batchelor, F. W. B. Einstein, and B. O. Patrick, *Inorg. Chem.* 40, 6026 (2001).
502. D. B. Leznoff, B.-Y. Xue, B. O. Patrick, V. Sanchez, and R. C. Thompson, *J. Chem. Soc., Chem. Commun.* 259 (2001).
503. P. Pyykkö, J. Li, and N. Runeberg, *Chem. Phys. Lett.* 218, 133 (1994).
504. B. L. Frankamp, A. K. Boal, and V. M. Rotello, *J. Am. Chem. Soc.* 124, 15146 (2002).
505. M. Laing, *Educ. Chem.* 160 (1993).

506. L. C. Allen and J. K. Burdett, *Angew. Chem. Int. Ed. Engl.* 34, 2003 (1995).
507. H. Burtscher and A. Schmidt-Ott, *Phys. Rev. Lett.* 48, 1734 (1982).
508. G. Boche, F. Basold, M. Marsch, and K. Harms, *Angew. Chem. Int. Ed. Engl.* 37, 1684 (1998).
509. T. A. Taton, C. A. Mirkin, and R. L. Letsinger, *Science* 289, 1757 (2000).
510. B. Dubertret, M. Calame, and A. J. Libchaber, *Nat. Biotechnol.* 19, 365 (2001).
511. "Colloid Gold: Principles, Methods and Applications." Academic Press, San Diego, 1989.
512. J. M. Thomas, B. F. G. Johnson, R. Raja, G. Sankar, and P. A. Midgley, *Acc. Chem. Res.* 36, 20 (2003).
513. J. Schulz, A. Roucoux, and H. Patin, *Chem. Rev.* 102, 3757 (2002).
514. B. Wang, S. Yin, G. Wang, A. Buldum, and J. Zhao, *Phys. Rev. Lett.* 86, 2046 (2001).
515. J. Nakamura, N. Kobayashi, S. Watanabe, and M. Aono, *Surf. Sci.* 482, 1266 (2001).
516. W. T. Geng and K. S. Kim, *Phys. Rev. B* 67, 233403 (2003).
517. J. Zhao, B. Calin, J. Han, and J. P. Lu, *Nanotechnology* 14, 501 (2003).
518. S. B. Suh, B. H. Hong, P. Tarakeshwar, S. J. Youn, S. Jeong, and K. S. Kim, *Phys. Rev. B* 67, 241402 (R) (2003).
519. M. Irie, T. Fukaminato, T. Sasaki, N. Tamai, and T. Kawai, *Nature* 420, 759 (2002).
520. J. Lehmann, S. Camalet, S. Kohler, and P. Hänggi, *Chem. Phys. Lett.* 368, 282 (2003).
521. T. Steenken, J. P. Telo, H. M. Novais, and L. P. Candeias, *J. Am. Chem. Soc.* 114, 4701 (1992).
522. D. Becker and M. D. Sevilla, "Advances in Radiation Biology." Academic Press, New York, 1993.
523. S. O. Kelly and J. K. Barton, *Science* 283, 375 (1999).
524. M. Ratner, *Nature* 397, 480 (1999).
525. N. A. Richardson, S. S. Wesolowski, and H. F. Schaefer, *J. Am. Chem. Soc.* 124, 10163 (2002).
526. J. Kim, J. M. Park, K. S. Oh, J. Y. Lee, S. Lee, and K. S. Kim, *J. Chem. Phys.* 106, 10207 (1997).
527. S. Lee, J. Kim, S. J. Lee, and K. S. Kim, *Phys. Rev. Lett.* 79, 2038 (1997).
528. J. Kim, S. B. Suh, and K. S. Kim, *J. Chem. Phys.* 111, 10077 (1999).
529. J. Kim, J. Y. Lee, K. S. Oh, J. M. Park, S. Lee, and K. S. Kim, *Phys. Rev. A* 59, R930 (1999).
530. S. B. Suh, H. M. Lee, J. Kim, J. Y. Lee, and K. S. Kim, *J. Chem. Phys.* 113, 5273 (2000).
531. H. M. Lee and K. S. Kim, *J. Chem. Phys.* 117, 706 (2002).
532. "Hydrated Electron" (E. J. Hart and M. Anber, Eds.). Wiley, New York, 1970.
533. M. A. Johnson and T. D. Mark, *Int. J. Mass Spectrom.* 220, 97 (2002).
534. K. H. Bowen and H. Haberland, in "Solvated Electron Clusters" (H. Haberland, Ed.). Springer-Verlag, Berlin, 1995.
535. "Electrons in Fluids" (J. Jortner and N. R. Kestner, Eds.). Springer-Verlag, New York, 1973.
536. C. C. Bradley, W. R. Anderson, J. J. McClelland, and R. J. Celotta, *Appl. Surf. Sci.* 141, 210 (1999).
537. V. Balzani, P. Ceroni, S. Gestermann, M. Gorka, C. Kaufmann, and F. Vögtle, *Tetrahedron* 58, 629 (2002).
538. H. Feniri, B.-L. Deng, and A. E. Ribbe, *J. Am. Chem. Soc.* 124, 11064 (2002).
539. K. Rurack and U. Resch-Genger, *Chem. Soc. Rev.* 31, 116 (2002).
540. S. C. McCleskey, A. Metzger, C. S. Simmons, and E. V. Anslyn, *Tetrahedron* 58, 621 (2002).
541. A. P. de Silva, H. Q. N. Gunaratne, T. Gunnlaugsson, A. J. M. Huxley, C. P. McCoy, J. T. Rademacher, and T. E. Rice, *Chem. Rev.* 97, 1515 (1997).
542. D. Majumdar, J. Kim, and K. S. Kim, *J. Chem. Phys.* 112, 101 (2000).
543. D. Majumdar, H. M. Lee, J. Kim, and K. S. Kim, *J. Chem. Phys.* 111, 5866 (1999).
544. S. G. Johnson and J. D. Joannopoulos, "Photonic Crystals: The Road from Theory to Practice." Kluwer, Boston, 2002.
545. E. Bakker and M. Telting-Diaz, *Anal. Chem.* 74, 2781 (2002).
546. R. Barrios, P. Skurski, and J. Simons, *J. Phys. Chem. B* 106, 7991 (2002).
547. C. J. Burrows and J. G. Muller, *Chem. Rev.* 98, 1109 (1998).
548. E. Fermi and E. Teller, *Phys. Rev.* 72, 406 (1947).
549. X. B. Wang, C. F. Ding, and L. S. Wang, *Phys. Rev. Lett.* 81, 3351 (1998).
550. L. S. Wang, C. F. Ding, X. B. Wang, and J. B. Nicholas, *Phys. Rev. Lett.* 81, 2667 (1998).
551. S. E. Bradforth and P. Jungwirth, *J. Phys. Chem. A* 106, 1286 (2002).
552. L. Lehr, M. T. Zanni, C. Frischkorn, R. Weinkauf, and D. M. Neumark, *Science* 284, 635 (1999).
553. W. Harneit1, C. Meyer, A. Weidinger, D. Suter, and J. Twamley, *Phys. Status. Solidi. B* 233, 453 (2002).
554. A. C. Grimsdale and K. Müllen, *Chem. Record* 1, 243 (2001).
555. G. R. Desiraju, *Acc. Chem. Res.* 29, 441 (1996).
556. V. Balzani and A. Credi, in "Molecular Level Devices in Supramolecular Science" (R. Ungaro and E. Dalcanale, Ed.), p. 1. Kluwer, Dordrecht, 1999.
557. H. Noji, R. Yasuda, M. Yoshida, and K. Kinosita, *Nature* 386, 299 (1997).
558. S. S. Smith, *Nano Lett.* 1, 51 (2001).
559. S. M. Dirk and J. M. Tour, *Tetrahedron* 59, 287 (2003).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanoscale Characterization of Biomaterials

E. Jallot

Université Blaise Pascal, 63177 Aubière Cedex, France

CONTENTS

1. INTRODUCTION

Biomaterials are defined as nonliving materials able to replace a part of the human body.

Life expectancy is now over 80 years. This increase in survivability, however, means that many people outlive the quality of their connective tissues, and the capacity of the human body to regenerate bony components that are lost or damaged is limited. Some 30 years ago, a revolution in medical care began with the successful replacement of tissues. Consequently, we have to find or to develop implants that might replace bony tissues. Two types of implants can be used: natural or synthetic [1].

The natural implants concern xenogenous, allogenous, and autogenous bone grafts. Xenogenous bone grafts are implants which come from another species. Allogenous bone grafts are implants which come from the same species. However, many problems were generally associated with them such as *in vivo* resorption, virus transfer, considerable care, high costs, and regular provocation of an immunological-defensive reaction, which limits their efficiency. Autogenous bone grafts are implants which come from the same body. These grafts are the most suitable because there is an excellent biocompatibility and no risk of virus transfer. However, removal of the bone grafts creates additional surgical trauma and its supply may not be available in sufficient quantity. To overcome all these problems, we have to find and to develop synthetic materials which might be used as bone substitutes or as prostheses [2].

The second possibility in the revolution to replace tissues was the development, or in many cases modification, of manmade materials to interface with living, host tissues. These synthetic implants are called biomaterials. The significant advantages of synthetic implants over natural implants are availability, reproducibility, and reliability. Good manufacturing practice, international standards, government regulations, and quality assurance testing minimize the risk of failure of implants. However, implants developed actually have serious disadvantages: problems of interfacial stability with host tissues, low mechanical properties, and production of wear particles. These problems limits biointegration of implants and their lifetimes. Efforts to improve properties of implants have to be made.

Actually, a lot of synthetic bone substitutes and prosthesis are available to repair bony tissues that are lost or damaged. The most widely used are polymers, metallic alloys (Ti_6Al_4V, Co–Cr, inox, ...), and bioceramics [alumina (Al_2O_3), zirconia (ZrO_2), calcium phosphates, bioactive glasses, biovitroceramics] [3, 4]. The ultimate goal of these materials is to reach full integration of the nonliving implant with living bone. With advances in ceramic technology, the application of calcium phosphate materials, bioactive glasses, and biovitroceramics as bone substitutes or as coatings on prostheses has received considerable attention, because they are highly biocompatible (well accepted in a biological environment) and they have bioactive properties [5–7]. These materials are capable, through physicochemical reactions, to establish a direct contact with bone [8]. However, many critical and complex reactions take place at the implant/bone tissue interface [9]. Structural and chemical evaluation of this interface is primordial to determine the success of an implant. Elemental composition and surface properties play a very important role in these reactions [10]. Knowledge of the elemental distribution at the biomaterial/bone tissue interface is important to understand the physicochemical mechanisms involved during the material integration and bone bonding [11, 12].

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (405–415)

2. ANALYSIS OF BIOMATERIAL INTERFACES

Structural and chemical evaluation of biomaterial/bone interfaces requires analysis at the nanometer level. Transmission electron microscopy (TEM), energy dispersive X-ray spectroscopy (EDXS), and electron energy loss spectroscopy (EELS) are methods which permit this analysis at the required resolution. TEM and associated techniques are powerful tools which can provide chemical and physical information: interlayer thickness, chemical species, local bonding, and the nature of crystalline or amorphous products [13, 14]. On the other hand, complementary techniques like particle induced X-ray emission (PIXE), secondary ion mass spectrometry (SIMS), and atomic force microscopy (AFM) are useful to better evaluate biomaterial/bone tissue interfaces.

2.1. TEM

Development of transmission electron microscopy permits one to obtain morphological and chemical information at the nanometer scale from the sample examined. Different types of interactions can occur between incident electrons and the sample. The majority of incident electrons are transmitted, elastically diffused, and inelastically diffused. Electrons are not diffused when incident electrons are transmitted without any interaction with the specimen. This phenomenon increases when the specimen thickness decreases. Electrons diffused elastically are electrons which are diffused by the electromagnetic field created by nucleus of atoms. During this interaction, they lose a very small amount of energy of the order of some eV. This diffusion is called elastic diffusion because there is no energy transferred. Electrons diffused inelastically lose energy during collisions with electrons from atoms of the sample. This energy loss is characteristic of the atoms with which this interaction occurs. Several other signals can be generated after electron beam–sample interactions: secondary electron emission, backscattering of incident electrons, cathodoluminescence, Auger electron emission, and X-ray emission. Secondary electrons are electrons ejected from the conduction band because of their low bonding energy. Their kinetic energy is of the order of 50 eV. These electrons give topographic information about the surface. Backscattered electrons are electrons which pass near nucleus of atoms and are backscattered by the electromagnetic field of the nucleus. Cathodoluminescence represents photons which have a wavelength between 0.4 and 0.8 μm (visible). These photons are emitted by materials like semiconductors, organic molecules during the irradiation by the electron beam. Moreover, the interaction between the high energy electrons from the beam and atoms from the sample lead to atom ionization. Electrons from the inner shells of atoms are ejected from their orbits, leaving the inner shells incompletely filled. This gap is filled by an electron from the outer shells. This transition can lead to an X-ray photon emission, with an energy equal to the energy difference between the two shells or to the emission of an electron from outer shells: Auger electron.

Morphological, structural information can be obtained by contrasted images of transmitted electrons [15]. Chemical information can be obtained from the electron beam–sample interaction [15, 16]. Two of the signals generated are the X-ray signal (which is used in EDXS) and the electron energy loss signal issued from electrons diffused inelastically (which is used in EELS). These two spectrometries are used to obtain elemental and eventually chemical information. EDXS is now well developed and allows the elemental analysis of all elements with $Z > 5$. This technique permits quantitative elemental analysis, concentration profiles, and elemental cartographies with a resolution of the order of 10 nm [17]. The minimal detectable concentration is of the order of 500 ppm. EDXS allows the elemental analysis of the biomaterial–tissue interface together with the study of its ultrastructure in the transmission electron microscope [18, 19].

2.2. EDXS

X-ray energy is characteristic of the element in which the electronic transition occurred and permits one to determine the elements present in the sample [20]. Atoms are composed with many shells and many transitions can take place [21]. The K line corresponds to an X-ray generated after the ionization of the K shell, the L line corresponds to an X-ray generated after the ionization of the L shell, and so on [22]. For example, a transition from L_{III} to K is designated $K_{\alpha 1}$ and from M_V to L_{III} is designated $L_{\alpha 1}$. Transitions do not occur with the same probability. The most intense in each series are the $\alpha 1$. Considering the same line ($K_{\alpha 1}$ for example) the X-ray energy increases with the atomic number. In an atom, the K lines have a lower energy than L lines and the L lines have a lower energy than the M lines.

On the other hand, the incident electron beam can be inelastically scattered by the electromagnetic field of the nucleus of atoms in the sample. The energy loss by the electron leads to an X-ray emission of corresponding energy which is called continuum or background [23]. This background does not permit one to determine the elements present in the sample. However, its intensity is proportional to the mass of the analyzed volume and can be used for quantitative calculations [24].

2.2.1. Instruments

A transmission electron microscope generates the electron beam and gives the ultrastructural information about the sample. The electron beam size is very important to obtain an optimal spatial resolution. Actually, the scanning transmission electron microscope (STEM) permits one to obtain a resolution under 1 nm. The X-rays generated in the specimen by the electron beam are collected by a semiconductor detector, amplified, and displayed with a multichannel analyzer associated with a computer. The spectrum is displayed and stored.

The type of detector usually used is a lithium drifted silicon semiconductor detector [in short Si(Li) detector] cooled to liquid nitrogen temperature. The resolution of a detector is defined as the full width half-maximal height of the Mn K_α peak which is at an energy of 5.9 keV [25]. Actually, resolution is of the order of 140 eV. This type of detector

is protected from contamination from the microscope by a beryllium entrance window of 8 μm thickness. However, this window absorbs low energy X-rays which limits the detection to elements heavier than Na ($Z > 11$). Recent development of these detectors with an ultrathin window or window less permits the analysis of all elements down to boron.

The spectrum is composed with characteristic X-ray peaks for various elements and backgrounds (Fig. 1). We can observe the presence of carbon (C K_α), oxygen (O K_α), copper (Cu L_α, Cu K_α, Cu K_β), sodium (Na K_α), silicon (Si K_α), phosphorus (P K_α), and calcium (Ca K_α, Ca K_β). Spectral lines like $\alpha 1$ and $\alpha 2$ (respectively $\beta 1$ and $\beta 2$) are so close to each other that they cannot be resolved by the spectrometer, and the convoluted line is simply denoted α (respectively β).

2.2.2. Quantitative Analysis

Calculations of elemental concentrations are made in the computer. The quantitative analysis of a specimen links the characteristic X-ray intensity to the number of atoms which emit this characteristic X-ray. Two methods can be used to calculate concentrations: the Hall method [26–28] and the Cliff and Lorimer method [29]. In the first method, Hall proposed an alternative approach in which the mass of the matrix is measured from the intensity of the background. The concentration of an element is proportional to the peak intensity/background intensity ratio. By using standards, this method permits one to quantify elemental concentrations without the detection of all elements. Concentrations are determined separately from each other. This method can be used with all types of detectors, even those with a beryllium entrance window. However, this method has a serious disadvantage. The background of the specimen is superposed to the background of the grid support. This contribution varies with the distance between the analyzed zone and the nearest bar of the grid. In order to minimize errors induced by this phenomenon, measurements must be done as far as possible from the bars of the grid. The second method does not use the background intensity. This method is based on the fact that the sum of all elemental concentrations (weight %) in the specimen is equal to unity. Then, the elemental concentration of each element can be calculated. However, this method needs the correct detection of all elements present in the sample and especially light elements which are highly present in biological specimens. Detectors with a beryllium window entrance cannot be used.

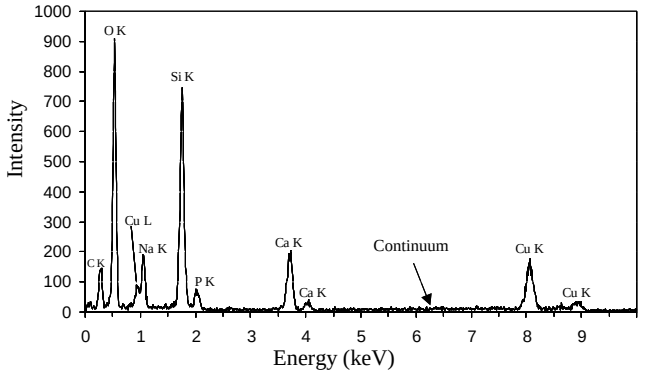


Figure 1. Spectrum showing characteristic peaks for various elements and the continuum (background).

2.2.3. Sample Preparation

For STEM and EDXS analysis at the nanometer scale the sample preparation is crucial. Several problems must be solved [30, 31].

- Ultrathin sections under 500 nm of the biomaterial–tissue interface must be prepared to perform the analysis at the resolution required. This is a very difficult step because the interface between hard materials and soft tissues must be in the sections.
- The specimen support should not contain elements of interest present in the studied sample.
- The preparation method must not change the chemical identity of the specimen. If the elements of interest in the sample are not firmly bonded, they can be solubilized during preparation [32].

Actually two preparation methods can be distinguished.

Conventional Preparation Method Biological ultrathin sections are prepared by a process including different steps [33]:

- fixation in aldehyde
- postfixation in osmium (osmium tetroxide, OsO_4)
- dehydration in increasing concentrations of alcohol
- embedding in resin
- sectioning with an ultramicrotome
- section staining with uranyl acetate and lead citrate
- carbon (C) or gold sputtering (Au)

This method allows a very good identification of the ultrastructure at the biomaterial–tissue interface [34, 35]. However, this preparation has disadvantages. Only firmly bonded elements can be analyzed. During fixation, the diffusible ions (Na, K, Cl, Mg, Ca) present in tissues and materials are lost from the sample [36]. During preparation, use of osmium, uranyl acetate, and lead citrate adds Os, U, and Pb to the sample. X-rays of these elements may interfere with X-rays of elements of interest in the sample. The gold sputtered may interfere too. Finally, chloride can be added to the sample by the resin. Table 1 summarizes different

Table 1. Common extraneous peaks causing artifacts during analysis.

Extraneous peak	Source	Interferes with these elements
Al K_α	specimen holder	Br L_α
Si K_α	contamination from microscope	Sr L_α
Cu K_α	grid	Zn K_α, Os L_α
Cu L_α	grid	Na
Pd L_α	coating	Cl K_α
Pd L_β	coating	K K_α
Os L_α	postfixation	Zn K_α
Os M_α	postfixation	P K_α
Au M_α	coating	P K_α, S K_α
Pb M_α	section staining	S K_α, Cl K_α
U M_α	section staining	K K_α

types of interference which can be generated by the sample preparation. For example, copper found in the spectrum in Figure 1 comes from the grid and copper is not present in the specimen.

Cryopreparation Methods Various methods for cryopreparation are available. The first step consists of a rapid freezing of the specimen by liquid ethane cooled by liquid nitrogen (−196 °C) or a metal cooled by liquid nitrogen [37, 38]. The high speed reduces considerably the formation of ice crystals. Then, cryosections are made with an ultracryomicrotome at a temperature of the order of −140 °C. Then, two types of analysis are possible:

- The sections are transferred to the microscope in the frozen-hydrated state and they can be studied in this state [39–43].
- The sections can be freeze-dried in the microscope chamber or they can be freeze dried and transferred to the microscope at room temperature for the analysis [38].

Other techniques including freeze-drying, freeze substitution then infiltration by resin, and finally sectioning can be used [44, 45].

These methods permit one to retain all elements of interest at their *in vivo* or *in vitro* position [46, 47]. They allow one to study ion distribution and ion diffusion in biomaterials during their interaction with tissues or in cells in contact with biomaterials [48]. Analysis is performed in a state near the *in vivo* or *in vitro* state.

2.3. EELS

EDXS is very useful for elemental analysis, but this technique is limited with its spatial resolution of analysis and its sensitivity. Electron energy loss spectroscopy appears to be a good complementary technique [49]. Its spatial resolution is of the order of the nanometer and the minimal detectable concentration is of the order of 400–500 ppm. During ionization of atoms present in the sample, electrons loss a certain amount of energy which is specific to the atom that is ionized. Detection of energy loss by the electrons that have passed through the sample permits one to determine the elemental composition and even fine structures. Sensitivity of this method for light elements is better than that of EDXS. Unfortunately, the sample must be very thin (less than 100 nm). The use of EELS is very limited for the study of biomaterials due to the difficulty in preparing sufficiently thin sections.

2.4. Selected-Area Electron Diffraction and High Resolution Transmission Electron Microscopy

High resolution transmission electron microscopy (HRTEM) and the electron diffraction mode, including selected-area electron diffraction (SAED) and microdiffraction, permit one to determine the crystalline character of a part of the studied specimens [50]. Electron beams of some nanometers in diameter are used to determine the crystalline structure. Materials are assessed as "amorphous" when electron diffraction patterns show broad and diffuse ring patterns and when there is no visible speckle in the rings. Crystalline materials give spotty diffraction patterns or spotty ring patterns. Finally, pattern analysis and calculations allow one to determine distances between atomic plans and the crystalline structure at the nanometer scale.

2.5. PIXE

Protons are charged particles which can generate characteristic X-rays in the same way that electrons do. The particle induced X-ray emission method used a 3 MeV energy proton beam to ionize atoms. This technique is used for trace element analysis because the peak to background ratio is much better than in EDXS [51]. The limit of detection of PIXE is several ppm. A further very important advantage is that PIXE can be used as a microbeam technique, which allows elemental mapping with a good spatial resolution. However, a proton beam cannot be as finely focused as an electron beam. Recent instrumental developments permit one to obtain proton beams in the order of 200–400 nm. This method is interesting to study trace elements locally at the biomaterials interface.

2.6. SIMS

Secondary ion mass spectrometry is used to analyze the surface composition and eventually a depth profile can be obtained because several layers are sputtered away. An ion beam is used to generate secondary ions from the surface specimen. The sensitivity of this method is of the order of 1 ppm and its resolution is near 100 nm.

2.7. AFM

During the last decade, developments in atomic force microscopy permit one to characterize the local properties of covered surfaces, namely the degree of coverage, the thickness of the layer, and the shape [52]. Interactions between a tip (cantilever) and specimen surface permit one to obtain images of surface structures at the nanometer scale. AFM has become a common experimental method in the investigation of biomaterials during the past 10 years [53], even if it is a delicate technique for measurements of soft materials as well as for evaluation and interpretation. This method allows one to determine adsorption of proteins on materials, surface rugosities induced by protein adsorption. Concerning bioactive materials, AFM is used to study changes of surface morphology and roughness during the bioactivity process. The structure of the apatite layer and the size of apatite crystals can be studied during interactions between bioactive materials and biological fluids [54, 55].

3. PHYSICOCHEMICAL REACTIONS AT BIOMATERIAL INTERFACES

3.1. Biomaterials Used to Replace Bony Tissues

Actually, various types of biomaterials are used in orthopaedic and maxillofacial surgery. These materials can generate different reactions at the implant/bone interface

and can be separated into three categories: biotolerant materials, bioinert materials, and bioactive materials [56–59].

3.1.1. Biotolerant Materials

Biotolerant materials are materials which are incorporated in bony tissues but with interposition of a layer of fibrous tissues between the materials and bone. For example, polymethyl methacrylate and cobalt–chrome (Co–Cr) alloy. The distance between the materials and bone is in the order of 1–20 μm [60]. The presence of the fibrous layer can lead to micromovements of the implant, corrosion, and the release of wear debris, which limits long-term stability.

3.1.2. Bioinert Materials

Bioinert materials are materials which are in contact with bone without leading to low tissue reactions and formation of a fibrous layer, for example: alumina (Al_2O_3), zirconia (ZrO_2), and titanium alloys [60]. However, there is no intimate or chemical contact between bone and implants. These materials are poorly degraded by the surrounding tissues and well accepted by the biological environment. Unfortunately, sometimes wear debris is released, which limits the efficiency of the implant and its long-term stability.

3.1.3. Bioactive Materials

Bioactive materials are defined by Hench [61] as materials capable of creating a chemical bond with surrounding tissues without interposition of a fibrous layer. This phenomenon leads to an intimate link between bone and materials. Two kinds of bioactive materials can be defined: resorbable materials and surface reactive materials.

After implantation in bony tissues, the material is progressively resorbed by bone cells (macrophages, osteoclasts), degraded by a chemical dissolution, and finally replaced by bone [62]. Material resorption and bone formation occur simultaneously in order to have a stability of the bone/material interface [63]. The fixation of these implants is called "biological fixation" with a porous ingrowth. These materials exhibit osteoconductive properties, defined as the characteristic of bone growth in porosities and bonding along the surface. Having the same chemical composition as natural bone, the calcium phosphate ceramics [64] (Table 2) and calcium carbonates ($CaCO_3$, coral) [65] are resorbable.

The dissolution process is highly dependent on the porosity, the crystalline structure, and the degree of crystallinity [66]. Calcium phosphate ceramics with macroporosities (300–400 μm) are more resorbable than calcium phosphate ceramics with microporosities (<10 μm) [67, 68]. A high crystallinity degree reduces the resorbability of calcium phosphates [69]. On the other hand, it is interesting to note that some polymers are resorbable. However, mechanisms of their degradation/resorption and biological effects of their degradation products are not well known.

Table 2. Some of the available calcium phosphates.

Calcium phosphates (synthetic)	Chemical formulae	Ca/P
Brushite, dicalcium phosphate dihydrate	$CaHPO_4 \cdot H_2O$	1
Monetite, dicalcium phosphate anhydrous	$CaHPO_4$	1
Octacalcium phosphate	$Ca_8H_2(PO_4)_6 \cdot 5H_2O$	1.33
Tricalcium phosphate	$Ca_3(PO_4)_2$	1.5
Calcium hydroxyapatite	$Ca_{10}(PO_4)_6(OH)_2$	1.67
Calcium fluorapatite	$Ca_{10}(PO_4)_6F_2$	1.67
Tetracalcium phosphate	$Ca_4P_2O_9$	2

During the last decade, considerable attention has been directed toward use of bioactive fixation of implants. The most bioactive materials are bioactive glasses, biovitroceramics, hydroxyapatites [$Ca_{10}(PO_4)_6(OH)_2$], and composite materials (hydroxyapatite–tricalcium phosphate, ...). At the interface between these materials and bone tissues a direct chemical bond is formed through a biological active apatite (Ca, P) layer [70–72]. This chemical reaction is the bioactivity property. An important advantage of bioactive fixation is that a bioactive bond to bone has a strength equal or greater than bone after 3 months. A bioactive bond to collagen fibrils of soft connective tissues is stronger than the cohesive bond of collagen fibrils. This high strength of both hard and soft tissues bonding to bioactive implants comes from the *in vivo* growth of an apatite layer of nanometer scale which binds to collagen fibrils [8].

3.2. Bioactivity Process

The bioactivity process differs with the type of material, its composition, and its structure. We can distinguish four types of bioactivity process.

3.2.1. Bioactivity Process of Bioactive Glasses

In 1970, Hench [10, 73] discovers bioactive glasses. They are amorphous materials with low mechanical properties which reduces their applications to prosthetic coatings [74, 75] and to fill bony defects [76–79]. The particular composition which initiates bioactivity is based on four oxides: (45%) SiO_2, (24.5%) CaO, (24.5%) Na_2O, and (6%) P_2O_5. The bioactivity properties of these materials depend on the percentage of these three oxides: SiO_2, CaO, Na_2O. An increase of SiO_2 considerably reduces the bioactivity and a high increase leads to a bioinert material [61]. An increase of CaO or Na_2O leads to a nonglass material. In fact, a very limited range of bioactive glass compositions, containing SiO_2–CaO–Na_2O–P_2O_5, that have less than 55% SiO_2 exhibit a high bioactivity and bond to both bone and soft connective tissues [80, 81]. Bioactive glasses can be obtained by melting the components at 1350 °C or by the sol–gel method at a lower temperature [82]. The synthesis of bioactive glasses by the sol–gel process was proposed in the last decade [83, 84]. The solutions to obtain bioactive gel-glasses were prepared from stoichiometric amounts of tetraethoxysilane, triethyl phosphate, and $Ca(NO_3)_2 \cdot 4H_2O$. Hydrolysis and condensation at low temperature create a highly interconnected three-dimensional gel network composed of $(SiO_4)^{4-}$ tetrahedral via bridging oxygen bonds or by Si–O–Ca or Si–O–P nonbridging bonds [85–88].

The bioactivity process has been studied *in vitro* during interactions between bioactive glasses and biological fluids (Table 3) or *in vivo* during interactions with bony tissues

Table 3. Ion concentrations (mmol dm^{-3}) of biological solutions used to study physicochemical reactions of bioactive materials.

	Na^+	K^+	Mg^{2+}	Ca^{2+}	Cl^-	HCO_3^-	HPO_4^{2-}	SO_4^{2-}
Simulated body fluid	142	5	1.5	2.5	147.8	4.2	1	0.5
Human plasma	142	5	1.5	2.5	103	27	1	0.5

[12, 89–92]. This process is decomposed in a complex series of physicochemical reactions including dissolution, diffusion, ionic exchange, and precipitation (Figure 2) [93–96]:

- Rapid exchange of alkali ions (Na^+) with H^+ or H_3O^+ from surrounding fluids through an exchange layer of the order of 200 nm [14].
- Loss of soluble Si to the solution resulting from breaking of Si–O–Si bonds. Soluble silicon migrates toward the surface and there is formation of Si–OH and Si–$(OH)_4$ groups at the surface. The exchange process consists of a flow of H_3O^+ ions and an equivalent flow of sodium ions into the solution from the glass represented as

$$SiO^-Na^+ + H_3O^+ \rightarrow SiOH + H_2O + Na^+$$

  Then, there is condensation and repolymerization of a hydrated Si rich layer on the surface depleted in alkali and Ca, P elements. This layer is a vitreous gel [54].
- Migration of Ca and P from the glass to the surface of the Si layer. Several authors described this phenomenon as a diffusion process through the Si rich layer.
- Growth of the vitreous gel by alkali exchange.
- Precipitation of an apatite (Ca–P) rich layer on top of the Si layer.
- Growth of the Ca–P rich layer. At the beginning the apatite crystals size is of the order of 200 nm and increases to some micrometers by consuming Ca^{2+}, PO_4^{3-}, CO_3^-, and HPO_4^{2-} from surrounding fluids. This layer incorporates collagen fibrils from surrounding tissues.

The formation of this apatite layer represents the bioactivity properties and permits a chemical link between the materials and the newly formed bone. The apatite crystals of the Ca–P layer and the apatite crystals of bone are intermingled each other which leads to a strong bonding [97]. Concerning bioactive gel-glasses, the 1–10 nm scale solid network of the gel is completely interpenetrated by pore liquid. The pore liquid consists of a highly structured hydrated layer with hydrated connective tissues. Biological molecules can exchange with these hydrated layers inside the pores and maintain their structure together with their biological properties [98, 99]. The nanometer sized pores of the gel-glass are proposed to act as initiation sites for apatite nucleation [100].

STEM is able to show unambiguously the presence or absence of thin amorphous or crystalline films at interfaces. Figure 3 shows bioactive glass particles cultured with bone cells (osteoblasts) for two days. Interactions between bioactive glass particles and the culture medium lead to the material dissolution, the formation of the Si layer, and finally the precipitation of an apatite layer. On the image, this layer appears as an electron dense layer. At this short time period, the apatite layer is of the order of 200 nm in thickness and its size will increases with time. In the case of a crystallized apatite layer, a needle-shaped structure will appear. The EDXS spectrum on this layer demonstrates the presence of Ca and P elements [75]. The apatite layer acts as a template for bone cell adhesion and differentiation [91]. Figure 4 shows a bioactive glass coating/bone interface at 3 months. We clearly notice the presence of the bioactive

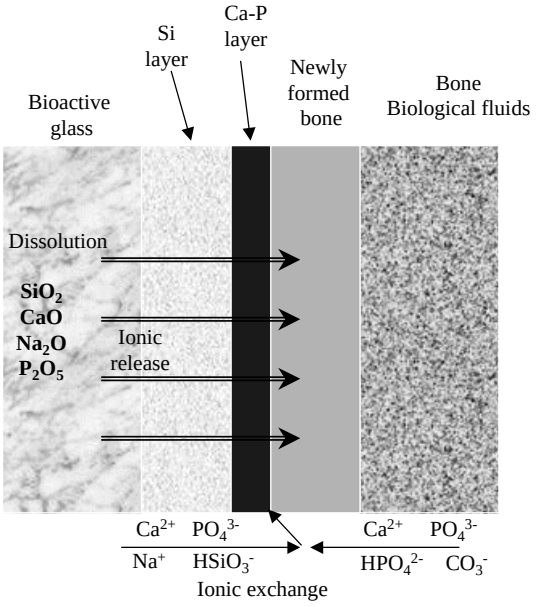


Figure 2. Schematic view of the bioactive glass/bone interface during the bioactivity process.

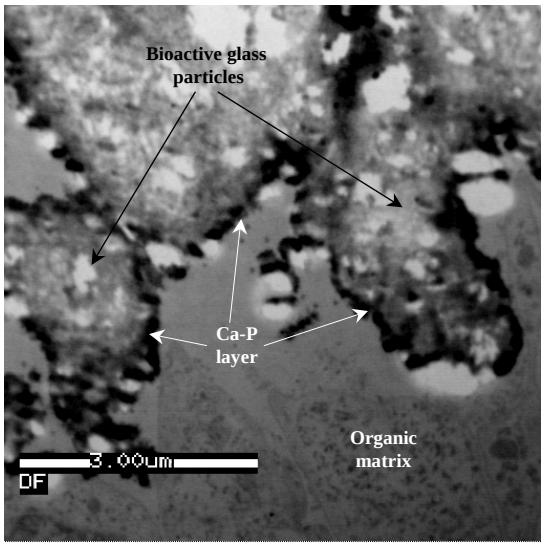


Figure 3. STEM micrograph of bioactive glass particles immersed in a biological solution with bone cells (osteoblasts) during two days. The particles are in dissolution and an electron dense layer appears (apatite layer). These particles are engulfed in an organic matrix.

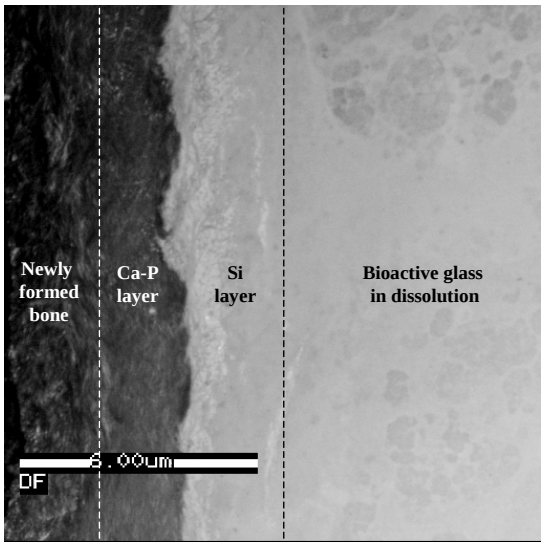


Figure 4. STEM micrograph of a bioactive glass coating/bone interface after 3 months of implantation. The bioactive glass coating is in dissolution and firmly fixed to newly formed bone through an apatite layer on top of a pure Si layer.

glass coating in dissolution, the Si layer, the apatite layer, and the newly formed bone. Intimate contact is observed between the glass and bone. Elemental profiles performed with EDXS permit one to determine the exact composition of the different zones and their evolution with time.

One of the main problems associated with the use of bioactive glasses as prosthetic coatings is high solubility, which might reduce the long-term stability of the glass. Extensive dissolution of the glass networks occurs within days. Optimization of bioactive glass properties concerns the compromise between bioactivity and solubility which is related to individual components [101–104]. For example, addition of Al_2O_3 can be used to control the solubility of the glass. However, this addition may inhibit the bone bonding [105]. Greenspan and Hench demonstrate that a concentration of Al_2O_3 higher than 2% inhibits bioactivity [106].

3.2.2. Bioactivity Process of Biovitroceramics

An *in vitro* ceramic is obtained by a temperature treatment of a glass in which nucleation factors are added in order to induce a partial or total crystallization of this glass. This treatment increases mechanical properties of the glass. A biovitroceramics has a structure and a particular chemical composition which are crucial for the bioactivity properties. Three types of biovitroceramics are elaborated [107, 108]:

- biovitroceramics composed of a vitreous matrix [(16.6%) MgO, (24.2%) CaO, (59.2%) SiO_2] in which apatite crystals ($Ca_{10}(PO_4)_6(O,F_2)$) are incorporated
- biovitroceramics (A–W) composed of a vitreous matrix in which apatite crystals and β-wollastonite crystals ($CaSiO_3$) are incorporated
- biovitroceramics composed of a vitreous matrix in which apatite crystals, β-wollastonite crystals, and whitelockite crystals ($3CaOP_2O_5$) are incorporated

All these biovitroceramics contain apatite crystals which are important in the bioactivity process. Studies of the bioactivity were made *in vitro* during interactions between biovitroceramics and biological fluids, and *in vivo* during interactions between biovitroceramics and bone [108, 109]. The bioactivity is based on (Fig. 5):

- dissolution of the vitreous matrix, wollastonite crystals, and whitelockite crystals (if they are incorporated)
- ionic release of $HSiO_3^-$, Ca^{2+}, Mg^{2+} at the surface
- precipitation of an apatite layer by consuming Ca^{2+}, PO_4^{3-}, CO_3^-, and HPO_4^{2-} from biological fluids

The formation of this apatite layer at the surface of biovitroceramics occurs without the formation of a Si rich layer [110, 111]. This layer of some micrometers in thickness permits a chemical bond between the materials and the newly formed bone.

3.2.3. Bioactivity Process of Hydroxyapatites

Synthetic hydroxyapatites are calcium phosphate ceramics elaborated under pressure and under temperature treatment of $Ca(NO_3)_2$, H_3PO_4, NH_4OH, and H_2O [112]. This ceramic is used in different forms: powders, bulk, or as coatings. For example, Figure 6 shows a STEM micrograph of nanostructures of hydroxyapatite powders. For bulk materials, various porosities are available [113]. We can distinguish microporosities (<10 μm) which permit diffusion of ions and fluids from macroporosities (100–600 μm) which permit cellular colonization. The bioactivity process was studied *in vitro* and *in vivo* [114–116]. It occurs under an acidic attack with H^+ at the material surface (Figure 7) [117]. This leads to the dissolution of hydroxyapatite crystals and a high release of Ca^{2+} and PO_4^{3-} [118, 119]. Concentration of calcium and phosphorus increase in the surrounding fluids and this supersaturation induces reprecipitation of apatite crystals at the ceramic

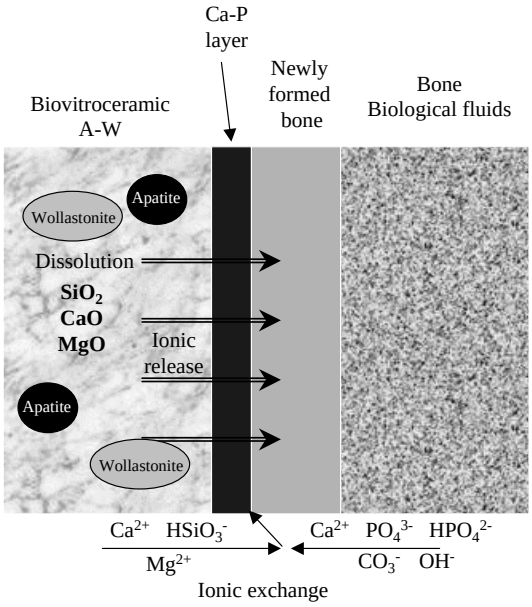


Figure 5. Schematic view of the biovitroceramic/bone interface during the bioactivity process.

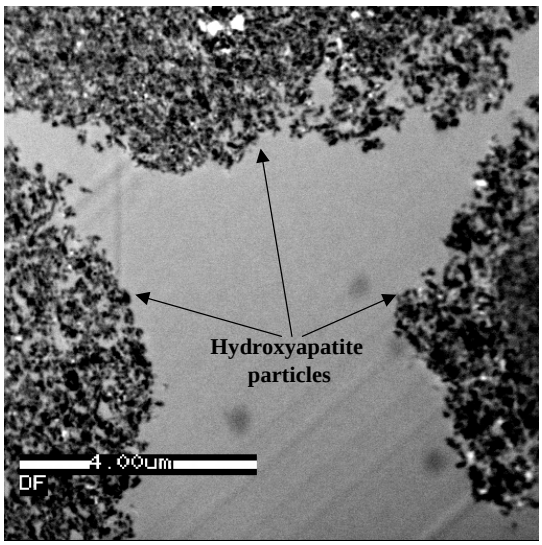


Figure 6. TEM micrograph showing nanostructures of hydroxyapatite particles.

surface [120]. These apatite crystals may incorporate Ca^{2+}, Mg^{2+}, CO_3^-, PO_4^{3-}, and organic molecules present in the surrounding fluids [121, 122]. This dissolution–reprecipitation process leads to the formation of a carbonated apatite layer at the material surface and permits a chemical bond with newly formed bone [114]. This layer is on the order of 200–800 nm [123–126]. Solubility of hydroxyapatites varies with different factors: porosity, grain size, crystallinity, sintering temperature [127–130]. An increase of the sintering temperature leads to an increase of the hydroxyapatite crystal size and finally reduces its solubility. For example, hydroxyapatite treated at 600 °C has crystals of 180 nm in size and crystals of 350 nm in size with a treatment at 1180 °C [129]. On the other hand, the solubility increases with the porosity and pore size.

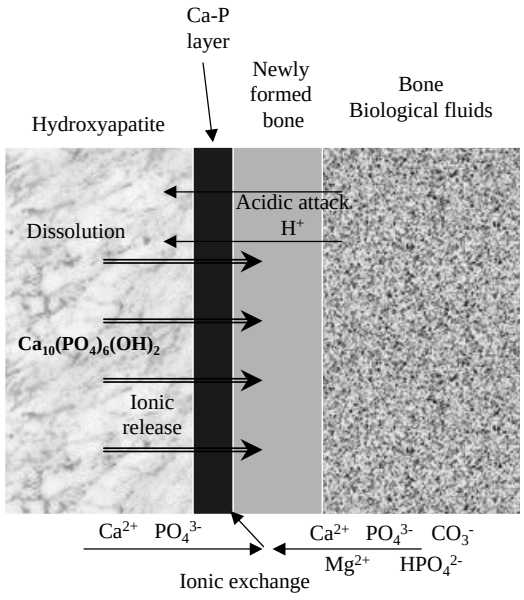


Figure 7. Schematic view of the hydroxyapatite/bone interface during the bioactivity process.

3.2.4. Bioactive Titanium

Recently, it has been found that even pure titania hydrogel prepared by the sol–gel method induces apatite formation [72, 131, 132]. When titanium is treated by immersion in alkaline solution (NaOH), a hydrated titanium oxide gel layer containing alkali ions is formed on its surface. This gel layer is dehydrated and densified to form an amorphous alkali titanate ($TiO_2 + Na^+$) layer by heat treatment below 600 °C. When the pretreated titanium is exposed to biological fluids, the alkali ions are released from the amorphous alkali titanate layer and hydronium ions enter into the surface layer, resulting in the formation of a titanium oxide hydrogel layer. The released Na^+ ions increase the degree of supersaturation of the soaking solution with respect to apatite by increasing pH, and titanium oxide hydrogel induces apatite nucleation on the titanium surface [133].

3.2.5. Apatite Precipitation Process

Determination of crystal size, nucleation, and growth of apatite during interactions with surrounding fluids is important to better understand the dissolution–precipitation process during bioactivity [134]. The combination of HRTEM, SAED, and nanoprobe electron diffraction allows one to determine morphology and structure of apatite crystals size [135]. The crystal size varies from a few nanometers to a micrometer [136]. The increase of the apatite crystal size occurs as a result of lower supersaturation in the crystallization system because of the absence of convection in space and the strong reduction in the rate of nucleation [137].

In some cases, the apatite precipitates are amorphous. During apatite precipitation, some ions like Mg^{2+} and CO_3^- are able to enter the forming apatite nuclei and thus inhibit their evolution to tiny apatite crystals [138]. These phenomena can promote a greater dissolution of the newly formed apatite precipitates [139–141].

4. CONCLUSION

Structural and chemical evaluation of bioactive biomaterial/bone tissue interfaces requires analysis at the nanometer scale. Transmission electron microscopy and associated techniques enable the chemical and physical characterization of bioactive material/bone tissue interfaces, providing information on interlayer thickness, chemical species, local bonding, and the nanostructural features which give rise to the interfacial properties, thereby enabling a full understanding not only of biomaterials after processing but also after interactions with bone tissues [142].

Bioactive materials are capable of bonding directly to bone through an apatite layer. This intimate link results from simultaneous physicochemical and biological processes. The bioactive materials in contact with bone tissues undergo dissolution with leaching and precipitation of apatite crystals occurs together with adsorption of organic compounds [143, 144]. Apatite crystals grow by epitaxial growth [145]. Simultaneously, bone cells mineralize an extracellular matrix on the materials composed with apatite crystals. Apatite crystals issued from the materials and apatite crystals issued from mineralization are intermingled which lead to a strong bonding between implants and bone.

This process of bioactivity occurs more rapidly in bioactive glasses and in biovitroceramics than in hydroxyapatite. Actually, these bioactive materials are developed and used in orthopaedic and maxillofacial surgery.

GLOSSARY

Bioactivity Property of a materials to form a direct chemical bond with bone tissues.

REFERENCES

1. M. B. Habal and A. H. Reddi, "Bone Grafts and Bone Substitutes." Saunders, Philadelphia, 1992.
2. Y. Fujishiro, L. L. Hench, and H. Oonishi, *J. Mater. Sci. Mater. Med.* 8, 649 (1997).
3. L. L. Hench, *J. Am. Ceram. Soc.* 74, 1487 (1991).
4. E. Jallot, H. Benhayoune, G. Weber, G. Balossier, and P. Bonhomme, *J. Phys. D* 33, 321 (2000).
5. M. Jarcho, *Clin. Orthop. Relat. Res.* 157, 259 (1981).
6. R. Z. Legeros, *Adv. Dent. Res.* 2, 164 (1988).
7. C. Ohtsuki, T. Kokubo, and T. Yamamuro, *J. Am. Ceram. Soc.* 75, 2094 (1992).
8. L. L. Hench, *Biomaterials* 19, 1419 (1998).
9. T. Kitsugi, T. Nakamura, M. Oka, Y. Senaha, T. Goto, and T. Shibuya, *J. Biomed. Mater. Res.* 30, 261 (1996).
10. L. L. Hench, R. J. Splinter, T. K. Greenlee, and W. C. Allen, *J. Biomed. Mater. Res.* 5, 117 (1971).
11. A. M. Gatti, G. Valdre, and A. Tombesi, *J. Biomed. Mater. Res.* 31, 475 (1996).
12. E. Jallot, H. Benhayoune, L. Kilian, J. L. Irigaray, H. Oudadesse, G. Balossier, and P. Bonhomme, *Surf. Interf. Anal.* 29, 314 (2000).
13. M. Neo, T. Nakamura, T. Kikutani, K. Kawanabe, and T. Kokubo, *J. Biomed. Mater. Res.* 27, 999 (1993).
14. E. Jallot, H. Benhayoune, L. Kilian, Y. Josset, and G. Balossier, *Langmuir* 17, 4467 (2001).
15. C. C. Gray, J. N. Chapman, W. A. P. Nicholson, B. W. Robertson, and R. P. Ferrier, *X-ray Spectrom.* 12, 163 (1983).
16. J. N. Chapman, C. C. Gray, B. W. Robertson, and W. A. P. Nicholson, *X-ray Spectrom.* 12, 153 (1983).
17. J. I. Goldstein, J. L. Costley, G. W. Lorimer, and S. J. B. Reed, in "Scanning Electron Microscopy" (O. Johari, Ed.), Vol. 1, p. 135. IITRI, Chicago, 1977.
18. C. Otsuki, Y. Aoki, T. Kokubo, Y. Bando, M. Neo, and T. Nakamura, *J. Am. Ceram. Soc.* 103, 449 (1995).
19. R. Z. Legeros and G. Daculsi, "Handbook of Bioactive Ceramics," Vol. 2, p. 1. CRC Press, Boca Raton, 1990.
20. C. J. Powell, *Rev. Mod. Phys.* 48, 33 (1976).
21. K. F. J. Heinrich, C. E. Fiori, and R. L. Myklebust, *J. Appl. Phys.* 50, 5589 (1979).
22. A. Langenberg and J. Van Eck, *J. Phys. B* 12, 1331 (1979).
23. R. H. Pratt, H. K. Tseng, C. M. Lee, and L. Kissel, At. *Data Nucl. Data Tables* 20, 175 (1977).
24. T. A. Hall, in "Physical Techniques in Biological Research," 2nd ed. (G. Oster, Ed.), Ch. 3. Academic Press, New York, 1971.
25. P. J. Statham, *X-ray Spectrom*. 5, 16 (1976).
26. T. A. Hall, *J. Microsc.* 117, 145 (1979).
27. T. A. Hall, *Micron Microsc. Acta* 17, 91 (1986).
28. T. A. Hall, *Scanning Micros.* 3, 461 (1989).
29. G. Cliff and G. W. Lorimer, *J. Microsc.* 103, 203 (1975).
30. G. M. Roomans, *J. Electron Micros. Tech.* 9, 3 (1988).
31. K. Zierold, *J. Microsc.* 161, 357 (1991).
32. B. L. Gupta and T. A. Hall, *Federation Proc.* 38, 144 (1979).
33. G. Nicaise, I. Gillot, A. K. Julliard, E. Keicher, S. Blaineau, J. Amsellem, J. C. Meyran, M. L. Hernandez-Nicause, B. Ciapa, and C. Gleyzal, *Scanning Microsc.* 3, 199 (1989).
34. M. L. Hernandez-Nicaise and J. Amsellem, *J. Ultrastruct. Res.* 72, 151 (1980).
35. J. M. Sautier, D. Septier, F. J. G. Cuisinier, J. R. Nefussi, M. Oboeuf, J. C. Voegel, N. Forest, and M. Goldberg, *Cells Mater.* 4, 357 (1994).
36. J. A. Chandler and S. Battersby, in "Microbeam Analysis in Biology" (C. P. Lechene and R. R. Warner, Eds.), p. 457. Academic Press, New York, 1979.
37. S. Blaineau, A. K. Julliard, J. Amsellem, and G. Nicaise, *Histochemistry* 87, 545 (1987).
38. K. Zierold, H. Hentschel, F. Wehner, and A. Wessing, *Scanning Microsc. Suppl.* 8, 117 (1994).
39. A. T. Marshall, *Scanning Electron Microsc.* 2, 335 (1980).
40. A. T. Marshall, in "Cryotechniques in Biological Electron Microscopy" (R. A. Steinbrecht and K. Zierold, Eds.), p. 240. Springer-Verlag, Berlin, 1987.
41. A. T. Marshall and W. Xu, *J. Microsc.* 190, 305 (1998).
42. P. Echlin and S. E. Taylor, *J. Microsc.* 141, 329 (1986).
43. P. Echlin, in "Low-Temperature Microscopy and Analysis." Plenum Press, New York, 1992.
44. D. M. R. Harvey, *J. Microsc.* 127, 209 (1982).
45. D. M. R. Harvey, J. L. Hall, and T. J. Flowers, *J. Microsc.* 107, 189 (1976).
46. A. T. Marshall, *Scanning Mircrosc. Suppl.* 8, 187 (1994).
47. A. T. Marshall, *Scanning Electron Mircrosc.* 2, 327 (1981).
48. E. M. Nkamgueu, J. J. Adnet, J. Bernard, K. Zierold, L. Kilian, E. Jallot, H. Benhayoune, and P. Bonhomme, *J. Biomed. Mater. Res.* 52, 587 (2000).
49. R. D. Leapman and R. L. Ornberg, *Ultramicroscopy* 24, 251 (1988).
50. E. I. Suvorova and P. A. Buffat, *J. Microsc.* 196, 46 (1999).
51. E. Jallot, J. L. Irigaray, H. Oudadesse, V. Brun, G. Weber, and P. Frayssinet, *Nucl. Instrum. Methods B* 142, 156 (1998).
52. M. Collaud Coen, R. Lehmann, P. Gröning, M. Bielmann, C. Galli, and L. Schlapbach, *J. Colloid Interf. Sci.* 233, 180 (2001).
53. A. Ikai, *Surf. Sci. Rep.* 26, 261 (1996).
54. E. Saiz, M. Goldman, J. M. Gomez-Vega, A. P. Tomsia, G. W. Marshall, and S. J. Marshall, *Biomaterials* 23, 3749 (2002).
55. A. Itälä, E. C. Nordstöm, H. Ylänen, H. T. Aro, and M. Hupa, *J. Biomed. Mater. Res.* 56, 282 (2001).
56. L. L. Hench and J. A. Paschall, *J. Biomed. Mater. Res. Symp.* 4, 25 (1973).
57. M. Neo, S. Kotani, T. Nakamura, and T. Yamamuro, *J. Biomed. Mater. Res.* 26, 1419 (1992).
58. M. Neo, S. Kotani, Y. Fujita, T. Nakamura, and T. Yamamuro, *J. Biomed. Mater. Res.* 26, 255 (1992).
59. E. Jallot, J. L. Irigaray, G. Weber, and P. Frayssinet, *Surf. Interf. Anal.* 27, 648 (1999).
60. J. F. Osborn and H. Newesely, in "Dynamic Aspect of the Implant/Bone Interface" (G. Heimke, Ed.), p. 111. Cark Hansen Verlag, Germany, 1980.
61. L. L. Hench, in "Handbook of Bioactive Ceramics" (T. Yamamuro, L. L. Hench, and J. Wilson, Eds.), Vol. 1, p. 7. CRC Press, Boca Raton, 1990.
62. E. Jallot, J. L. Irigaray, H. Oudadesse, V. Brun, G. Weber, and P. Frayssinet, *Eur. Phys. J.* 6, 205 (1999).
63. G. Daculsi, R. Z. Legeros, E. Nery, and K. Lynch, *J. Biomed. Mater. Res.* 23, 883 (1988).
64. R. Z. Legeros, *Clin. Mater.* 14, 65 (1993).
65. C. J. Damien, J. L. Ricci, P. Christel, H. Alexander, and J. L. Patat, *Calcified Tissue Int.* 55, 151 (1994).
66. G. Daculsi, R. Z. Legeros, and D. Mitre, *Calcified Tissue Int.* 46, 20 (1990).
67. H. Benhayoune, E. Jallot, P. Laquerrière, G. Balossier, P. Bonhomme, and P. Frayssinet, *Biomaterials* 21, 235 (2000).
68. C. P. A. T. Klein, A. A. Driessen, K. De Groot, and A. Van Den Hoof, *J. Biomed. Mater. Res.* 17, 769 (1983).

69. C. P. A. T. Klein, A. A. Driessen, and K. De Groot, *Biomaterials* 5, 157 (1984).
70. K. J. J. Pajamaki, T. S. Lindholm, and O. H. Andersson, *J. Mater. Sci. Mater. Med.* 6, 14 (1995).
71. U. Gross and V. Strunz, *J. Biomed. Mater. Res. Symp.* 7, 503 (1985).
72. L. Jonasova, F. A. Müller, A. Helebrant, J. Strnad, and P. Greil, *Biomaterials* 23, 3095 (2002).
73. L. L. Hench, *Med. Instrum.* 57, 136 (1973).
74. R. J. Furlong and J. F. Osborn, *J. Bone Joint Surg.* 77, 534 (1995).
75. A. Merolli, A. Cacchioli, L. Giannotta, and P. Tranquilli Leali, *J. Mater. Sci. Mater. Med.* 12, 727 (2001).
76. E. Shepers, M. De Clerq, P. Ducheyne, and R. Kempeneers, *J. Oral Rehabil.* 18, 439 (1991).
77. L. L. Hench and J. K. West, *Life Chem. Rep.* 13, 187 (1996).
78. Ö. H. Andersson, K. H. Karlsonn, and K. Kangasniemi, *J. Non-cryst. Solids* 119, 290 (1990).
79. Ö. H. Andersson, G. Liu, K. Kangasniemi, and J. Juhanoja, *J. Mater. Sci. Mater. Med.* 3, 145 (1992).
80. W. Cao and L. L. Hench, *Ceramics Int.* 22, 493 (1996).
81. O. Peitl, E. D. Zanotto, and L. L. Hench, *J. Non-cryst. Solids* 292, 115 (2001).
82. W. Gong, A. Abdelouas, and W. Lutze, *J. Biomed. Mater. Res.* 54, 320 (2001).
83. R. Viitala, M. Jokinen, T. Peltola, K. Gunnelius, and J. B. Rosenholm, *Biomaterials* 23, 3073 (2002).
84. C. Kinowski, M. Bouazaoui, R. Bechara, L. L. Hench, J. M. Nedelec, and S. Turell, *J. Non-cryst. Solids* 291, 143 (2002).
85. L. L. Hench and R. Orefice, in "Kirk–Othmer Encyclopedia of Chemical Technology," 4th ed., Vol. 22, p. 497. Wiley, New York, 1997.
86. R. Li, A. E. Clark, and L. L. Hench, in "Chemical Processing of Advanced Materials" (L. L. Hench and J. K. West, Eds.), p. 627. Wiley, New York, 1992.
87. M. M. Pereira, A. E. Clark, and L. L. Hench, *J. Mater. Synth. Proc.* 2, 189 (1994).
88. M. M. Pereira, A. E. Clark, and L. L. Hench, *J. Biomed. Mater. Res.* 18, 693 (1994).
89. L. L. Hench and J. Wilson, *Science* 226, 630 (1984).
90. C. Y. Kim, A. E. Clark, and L. L. Hench, *J. Non-cryst. Solids* 113, 195 (1988).
91. C. Loty, J. M. Sautier, M. T. Tan, M. Oboeuf, E. Jallot, H. Boulekbache, D. Greenspan, and N. Forest, *J. Bone Miner. Res.* 16, 231 (2001).
92. A. El Ghannam, P. Ducheyne, and I. M. Shapiro, *J. Biomed. Mater. Res.* 29, 359 (1995).
93. I. D. Xynos, A. J. Edgar, L. D. K. Buttery, L. L. Hench, and J. M. Polak, *Biochem. Bioph. Res. Co.* 276, 461 (2000).
94. W. C. A. Vrouwenvelder, G. C. Groot, and K. De Groot, *Biomaterials* 13, 382 (1992).
95. A. El Ghannam, P. Ducheyne, and I. M. Shapiro, *Biomaterials* 18, 295 (1997).
96. E. Jallot, H. Benhayoune, L. Kilian, J. L. Irigaray, Y. Barbotteau, G. Balossier, and P. Bonhomme, *J. Colloid Interf. Sci.* 233, 83 (2001).
97. O. Pietrement and E. Jallot, *Nanotechnology* 13, 18 (2002).
98. D. Avnir, L. C. Klein, D. Levy, U. Schubert, and A. B. Wojcik, in "The Chemistry of Organosilicon Compounds," part 2 (Y. Apeloig and Z. Rappoport, Eds.), Wiley, Chichester, 1997.
99. J. Livage, C. Roux, J. M. Da Costa, I. Desportes, and J. E. Quinson, *J. Sol–Gel Sci. Technol.* 7, 45 (1996).
100. J. R. Jones and L. L. Hench, *Mater. Sci. Technol.* 17, 891 (2001).
101. C. Ohtsuki, T. Kokubo, and T. Yamamuro, *J. Non-cryst. Solids* 143, 84 (1992).
102. U. Gross and V. Strunz, *J. Biomed. Mater. Res. Symp.* 7, 503 (1985).
103. M. Vallet-Regi, I. Izquierdo-Barba, and A. J. Salinas, *J. Biomed. Mater. Res.* 46, 560 (1999).
104. M. Ogino, F. Ohuchi, and L. L. Hench, *J. Biomed. Mater. Res.* 14, 55 (1980).
105. K. Kangasniemi and A. Yli-Urpo, in "Handbook of Bioactive Ceramics" (J. Wilson, Ed.), Vol. 1, p. 97. CRC Press, Boston, 1990.
106. D. C. Greenspan and L. L. Hench, *J. Biomed. Mater. Res.* 10, 503 (1976).
107. T. Kokubo, H. Kushitani, C. Ohtsuki, S. Sakka, and T. Yamamuro, *J. Mater. Sci. Mater. Med.* 3, 79 (1992).
108. T. Kokubo, S. Ito, Z. T. Huang, T. Hayashi, S. Sakka, T. Kitsugi, and T. Yamamuro, *J. Biomed. Mater. Res.* 24, 331 (1990).
109. P. Li, C. Ohtsuki, T. Kokubo, K. Nakanishi, N. Soga, T. Nakamura, and T. Yamamuro, *J. Am. Ceram. Soc.* 75, 2094 (1992).
110. T. Kokubo, *Biomaterials* 12, 155 (1991).
111. T. Kokubo, *J. Ceram. Soc. Jpn.* 99, 965 (1991).
112. R. Aoba and E. C. Moreno, *J. Dent. Res.* 63, 874 (1984).
113. P. S. Eggli, W. Muller, and R. K. Schenk, *Clin. Orthopd. Relat. R.* 232, 127 (1988).
114. R. Z. Legeros and G. Daculsi, in "Handbook of Bioactive Ceramics" (T. Yamamuro, L. L. Hench, and J. Wilson, Eds.), Vol. 2, p. 17. CRC Press, Boca Raton, 1990.
115. I. Orly, M. Gregoire, M. Menanteau, and M. Dard, *Adv. Biomater.* 8, 211 (1988).
116. E. Jallot, *Med. Eng. Phys.* 20, 697 (1998).
117. G. Daculsi, G. Kerebel, and L. M. Kerebel, *Caries Res.* 13, 277 (1979).
118. G. Daculsi, R. Z. Legeros, and D. Mitre, *Calcified Tissue Int.* 45, 95 (1989).
119. R. Z. Legeros, I. Orly, M. Gregoire, and J. Kaimiroff, *J. Dent. Res.* 67, 177 (1988).
120. I. Orly, M. Gregoire, M. Menanteau, M. Heughebaert, and B. Kerebel, *Calcified Tissue Int.* 45, 20 (1989).
121. R. Z. Legeros, M. H. Taheri, G. M. Quirolgico, and J. P. Legeros, *Scanning Electron Microsc.* 407 (1983).
122. R. Z. Legeros, R. Kijkowska, T. Abergas, and J. P. Legeros, *J. Dent. Res.* 65, 293 (1986).
123. H. N. Denissen, K. De Groot, P. C. Makkes, A. Van Den Hoof, and P. J. Klopper, *J. Biomed. Mater. Res.* 14, 730 (1980).
124. J. M. Sautier, J. R. Nefussi, and N. Forest, *Cells Mater.* 1, 209 (1991).
125. J. M. Sautier, J. R. Nefussi, and N. Forest, *Biomaterials* 12, 400 (1992).
126. G. C. L. De Lange, C. De Putter, and F. L. J. A. De Wijs, *J. Biomed. Mater. Res.* 24, 829 (1990).
127. K. De Groot, *Ann. New York Acad. Sci.* 523, 227 (1988).
128. R. Z. Legeros and X. F. Chang, *J. Dent. Res.* 68, 215 (1989).
129. P. Laquerrière, L. Kilian, A. Bouchot, E. Jallot, A. Grandjean, M. Guenounou, G. Balossier, P. Frayssinet, and P. Bonhomme, *J. Biomed. Mater. Res.* 58, 238 (2001).
130. O. Malard, J. M. Bouler, J. Guicheux, D. Heymann, P. Pilet, C. Coquard, and G. Daculsi, *J. Biomed. Mater. Res.* 46, 103 (1999).
131. P. Li, C. Ohtsuki, T. Kokubo, K. Nakanishi, N. Soga, and K. De Groot, *J. Biomed. Mater. Res.* 28, 7 (1994).
132. P. Li, I. Kangasniemi, K. De Groot, and T. Kokubo, *J. Am. Ceram. Soc.* 5, 1307 (1994).
133. T. Kokubo, F. Miyaji, and H. M. Kim, *J. Am. Ceram. Soc.* 4, 1127 (1996).
134. G. Daculsi, R. Z. Legeros, J. P. Legeros, and D. Mitre, *J. Biomed. Mater. Res.* 2, 147 (1991).
135. E. I. Suvorova and P. A. Buffat, *J. Microsc.* 196, 46 (1999).
136. H. E. Lundager Madsen, F. Christensson, L. E. Polyak, E. I. Suvorova, M. O. Kliya, and A. A Chernov, *J. Cryst. Growth* 152, 191 (1995).
137. E. I. Suvorova, F. Christensson, H. E. Lundager Madsen, and A. A. Chernov, *J. Cryst. Growth* 186, 262 (1998).
138. G. Daculsi, R. Z. Legeros, M. Heughebaert, and I. Barbieux, *Calcified Tissue Int.* 46, 20 (1990).

139. E. Jallot, H. Benhayoune, L. Kilian, J. L. Irigaray, G. Balossier, and P. Bonhomme, *J. Phys. D* 33, 2775 (2000).
140. R. Z. Legeros, in "Tooth Enamel IV" (R. W. Fearnhead and S. Suga, Eds.), p. 32. Elsevier, Amsterdam, 1984.
141. R. Z. Legeros and R. Kijkowska, *J. Dent. Res.* 68, 1003 (1989).
142. E. Verne, C. Vitale Brovarone, C. Moisescu, E. Ghisolfi, and E. Marmo, *Acta Mater.* 48, 4667 (2000).
143. L. L. Hench, in "The Bone–Biomaterials Interface" (J. E. Davies, Ed.), p. 33. Univ. of Toronto Press, Toronto, 1991.
144. B. Kasemo and J. Lausmaa, in "The Bone–Biomaterials Interface" (J. E. Davies, Ed.), p. 19. Univ. of Toronto Press, Toronto, 1991.
145. R. Z. Legeros, I. Orly, M. Gregoire, and G. Daculsi, in "The Bone–Biomaterials Interface" (J. E. Davies, Ed.), p. 76. Univ. of Toronto Press, Toronto, 1991.

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanoscale Dilute Magnetic Semiconductors

S. J. Pearton, C. R. Abernathy

University of Florida, Gainesville, Florida, USA

Y. D. Park

Seoul National University, Seoul, Korea

CONTENTS

1. WHAT IS SPINTRONICS?

Two of the most successful technologies in existence today have created the Si integrated circuit (IC) industry and the data storage industry. Both continue to advance at a rapid pace. In the case of ICs, the number of transistors on a chip doubles about every 18 months according to Moore's law. For magnetic hard disk drive technology, a typical desk-top computer drive today has a 40 Gbyte/disk capacity, whereas in 1995 this capacity was ~1 Gbyte/disk. Since 1991, the overall bit density on a magnetic head has increased at an annual rate of 60–100% and is currently ~10.7 Gbits/in^2 [1]. The integrated circuits operate by controlling the flow of carriers through the semiconductor by applied electric fields. The key parameter therefore is the charge on the electrons or holes. For the case of magnetic data storage, the key parameter is the spin of the electron, as spin can be thought of as the fundamental origin of magnetic moment. The characteristics of ICs include high speed signal processing and excellent reliability, but the memory elements are volatile. The stored information is lost when the power is switched off, as data is stored as charge in capacitors (i.e., Dynamic Random Access Memories (DRAMs)). A key advantage of magnetic memory technologies is that they are nonvolatile since they employ ferromagnetic materials which by nature have remanence.

The emerging field of semiconductor spin transfer electronics (spintronics) seeks to exploit the spin of charge carriers in semiconductors. It is widely expected that new functionalities for electronics and photonics can be derived if the injection, transfer, and detection of carrier spin can be controlled above room temperature. Among this new class of devices are spin transistors operating at very low powers for mobile applications that rely on batteries, optical emitters with encoded information through their polarized light output, fast nonvolatile semiconductor memory, and integrated magnetic/electronic/photonic devices ("electromagnetism-on-a-chip"). Since the magnetic properties of ferromagnetic semiconductors are a function of carrier concentration in the material in many cases, it will be possible to have electrically or optically controlled magnetism through field-gating of transistor structures or optical excitation to alter the carrier density. This novel control of magnetism has already been achieved electronically and optically in an InMnAs metal–insulator semiconductor structure at low temperatures [2, 3] and electronically in Mn:Ge [4]. A number of recent reviews have covered the topics of spin injection, coherence length, and magnetic properties of materials systems such as (Ga,Mn)As [5–7], (In,Mn)As [5–7], and (Co,Ti)O_2 [8] and the general areas of spin injection from metals into semiconductors and applications of the spintronic phenomena [9–12]. The current interest in magnetic semiconductors can be traced to difficulties in

ISBN: 1-58883-063-2/$35.00

injecting spins from a ferromagnetic metal into a semiconductor [13, 14], an idea that can be traced to fruitful research in epitaxial preparation of ferromagnetic transitional metals on semiconductor substrates [15]. A theory first proposed by Schmidt et al. [16] points out that due to the dissimilar materials properties of a metal and semiconductor, an efficient spin injection in the diffusive transport regime is difficult unless the magnetic material is nearly 100% spin polarized, that is, half-metallic [17]. Although there have been recent reports of successful and efficient spin injection from a metal to a semiconductor even at room temperature by ballistic transport (i.e., Schottky barriers and tunneling) [18], the realization of functional spintronic devices requires materials with ferromagnetic ordering at operational temperatures compatible with existing semiconductor materials.

2. MATERIALS SELECTION

There are two major criteria for selecting the most promising materials for semiconductor spintronics. First, the ferromagnetism should be retained to practical temperatures (i.e., >300 K). Second, it would be a major advantage if there were already an existing technology base for the material in other applications. Most of the work in the past has focused on (Ga,Mn)As and (In,Mn)As. There are indeed major markets for their host materials in infrared (IR) light-emitting diodes and lasers and high speed digital electronics (GaAs) and magnetic sensors (InAs). Most of the past attention on ferromagnetic semiconductors focused on the (Ga,Mn)As [19–42] and (In,Mn)As [43–50] systems. In samples carefully grown single-phase by molecular beam epitaxy (MBE), the highest Curie temperatures reported are ~110 K for (Ga,Mn)As and ~35 K for (In,Mn)As. For ternary alloys such as $(In_{0.5}Ga_{0.5})_{0.93}Mn_{0.07}As$, the Curie temperature is also low, ~110 K [51]. A tremendous amount of research on these materials systems has led to some surprising results, such as the very long spin lifetimes and coherence times in GaAs [4] and the ability to achieve spin transfer through a heterointerface [52–69], either of semiconductor/semiconductor or metal semiconductor. One of the most effective methods for investigating spin-polarized transport is by monitoring the polarized electroluminescence output from a quantum-well light-emitting diode into which the spin current is injected. Quantum selection rules relating the initial carrier spin polarization and the subsequent polarized optical output can provide a quantitative measure of the injection efficiency [67, 69, 70].

There are a number of essential requirements for achieving practical spintronic devices in addition to the efficient electrical injection of spin-polarized carriers. These include the ability to transport the carriers with high transmission efficiency within the host semiconductor or conducting oxide, to detect or collect the spin-polarized carriers, and to control the transport through external means such as biasing of a gate contact on a transistor structure. The observation of spin current-induced switching in magnetic heterostructures is an important step in realizing practical devices [71]. Similarly, Nitta et al. [72] demonstrated that a spin–orbit interaction in a semiconductor quantum well could be controlled by applying a gate voltage. These key aspects of spin injection, spin-dependent transport, manipulation, and detection form the basis of current research and future technology. The use of read sensors based on metallic spin valves in disk drives for magnetic recording is already a $US100 B per year industry. It should also be pointed out that spintronic effects are inherently tied to nanotechnology, because of the short (~1 nm) characteristic length of some of the magnetic interactions. Combined with the expected low power capability of spintronic devices, this should lead to extremely high packing densities for memory elements. A recent review of electronic spin injection, spin transport, and spin detection technologies has recently been given by Buhrman [7], as part of a very detailed and comprehensive study of the status and trends of research into spin electronics in Japan, Europe, and the United States. The technical issues covered fabrication and characterization of magnetic nanostructures, magnetism and spin control in these structures, magneto-optical properties of semiconductors and magneto-electronics and devices. The nontechnical issues covered included industry and academic cooperation and long-term research challenges. The panel findings are posted on a web site [7].

In this chapter we focus on a particular and emerging aspect of spintronics, namely recent developments in achieving practical magnetic ordering temperatures in technologically useful semiconductors [73–79]. While the progress in synthesizing and controlling the magnetic properties of III-arsenide semiconductors has been astounding, the reported Curie temperatures are too low to have significant practical impact. A key development that focused attention on wide bandgap semiconductors as being the most promising for achieving high Curie temperatures was the work of Dietl et al. [80]. They employed the original Zener model of ferromagnetism [81] to predict T_C values exceeding room temperature for materials such as GaN and ZnO containing 5% of Mn and a high hole concentration (3.5×10^{20} cm^{-3}).

Other materials for which room temperature ferromagnetism has been reported include (Cd,Mn)GeP_2 [74], (Zn,Mn)GeP_2 [75], $ZnSnAs_2$ [76], (Zn,Co)O [77], and (Co,Ti)O_2 [8, 78] as well as Eu chalcogenides and others that have been studied in the past [79]. Some of these chalcopyrites and wide bandgap oxides have interesting optical properties, but they lack a technology and experience base as large as that of most semiconductors.

The key breakthrough that focused attention on wide bandgap semiconductors as being the most promising for achieving practical ordering temperatures was the theoretical work of Dietl et al. [80]. They predicted that cubic GaN doped with ~5 at.% of Mn and containing a high concentration of holes (3.5×10^{20} cm^{-3}) should exhibit a Curie temperature exceeding room temperature. In the period following the appearance of this work, there has been tremendous progress both on the realization of high-quality (Ga,Mn)N epitaxial layers and on the theory of ferromagnetism in these so-called dilute magnetic semiconductors (DMS). The term DMS refers to the fact that some fraction of the atoms in a nonmagnetic semiconductor like GaN is replaced by magnetic ions. A key, unanswered question is whether the resulting material is indeed an alloy of (Ga,Mn)N or

whether it remains as GaN with clusters, precipitates, or second phases that are responsible for the observed magnetic properties [82].

3. MECHANISMS OF FERROMAGNETISM

Figure 1 shows some of the operative mechanisms for magnetic ordering in DMS materials. Two basic approaches to understanding the magnetic properties of dilute magnetic semiconductors have emerged. The first class of approaches is based on mean field theory which originates in the original model of Zener [81]. The theories that fall into this general model implicitly assume that the dilute magnetic semiconductor is a more or less random alloy, for example, (Ga,Mn)N, in which Mn substitutes for one of the lattice constituents. The second class of approaches suggests that the magnetic atoms form small (a few atoms) clusters that produce the observed ferromagnetism [82]. A difficulty in experimentally verifying the mechanism responsible for the observed magnetic properties is that depending on the growth conditions employed for growing the DMS material, it is likely that one could readily produce samples that span the entire spectrum of possibilities from single-phase random alloys to nanoclusters of the magnetic atoms to precipitates and second phase formation. Therefore, it is necessary to decide on a case-by-case basis which mechanism is applicable. This can only be achieved by a careful correlation of the measured magnetic properties with materials analysis methods that are capable of detecting other phases or precipitates. If, for example, the magnetic behavior of the DMS is characteristic of that of a known ferromagnetic second phase [such as MnGa or Mn_4N in (Ga,Mn)N], then clearly the mean field models are not applicable. To date, most experimental reports concerning room temperature ferromagnetism in DMS employ X-ray diffraction, selected-area diffraction patterns, transmission electron microscopy, photoemission, or X-ray absorption [including extended X-ray absorption fine structure (EXAFS), as discussed later] to determine whether the magnetic atoms are substituting for one of the lattice constituents to form an alloy. Given the level of dilution of the magnetic atoms, it is often very difficult to categorically determine the origin of the ferromagnetism. Indirect means such as superconducting quantum interference device (SQUID) magnetometer measurements to exclude any ferromagnetic intermetallic compounds as the source of magnetic signals, and even the presence of what is called the anomalous or extraordinary Hall effect, that has been widely used to verify a single-phase system, may be by itself insufficient to characterize a DMS material. It could also certainly be the case that magnetically active clusters or second phases could be present in a pseudo-random alloy and therefore that several different mechanisms could contribute to the observed magnetic behavior. There is a major opportunity for the application of new, element- and lattice position-specific analysis techniques, such as the various scanning tunneling microscopies and Z-contrast scanning transmission electron microscopy among others for revealing a deeper microscopic understanding of this origin of ferromagnetism in the new DMS materials.

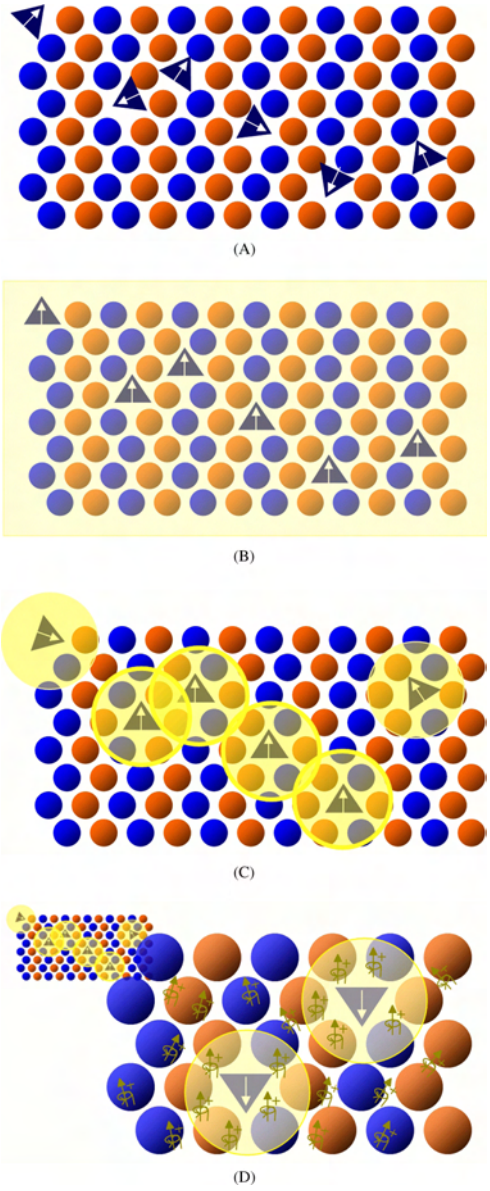


Figure 1. Semiconductor matrix with high concentrations of magnetic impurities (i.e., Mn), randomly distributed (defects), can be insulators (A) for group II–VI materials where divalent Mn ions occupy group II sites. At high concentrations, Mn ions are antiferromagnetically coupled, but at dilute limits, atomic distances between magnetic ions are large, and antiferromagnetic coupling is weak. For the cases where there is high concentrations of carriers (B) [i.e., (Ga,Mn)As where Mn ions behave as acceptors and provide magnetic moment as they occupy trivalent Ga sites], the carriers are thought to mediate ferromagnetic coupling between magnetic ions. Between near insulating and metallic cases, at low carrier regimes, hole carrier concentrations are localized near the magnetic impurity. Below certain temperatures, a percolation network (C) is formed in which clusters of holes are delocalized and hop from site to site, which energetically favors maintaining the carriers' spin orientation during the process, an effective mechanism for aligning Mn moments within the cluster network. Alternatively, at percolation limits, a localized hole near the magnetic impurity is polarized, and the energy of the system is lowered when the polarization of the localized holes are parallel (D).

The mean field approach basically assumes that the ferromagnetism occurs through interactions between the local moments of the Mn atoms, which are mediated by free holes in the material. The spin–spin coupling is also assumed to be a long-range interaction, allowing use of a mean field approximation [80, 83, 84]. In its basic form, this model employs a virtual-crystal approximation to calculate the effective spin density due to the Mn ion distribution. The direct Mn–Mn interactions are antiferromagnetic so that the Curie temperature, T_C, for a given material with a specific Mn concentration and hole density (derived from Mn acceptors and/or intentional shallow level acceptor doping), is determined by a competition between the ferromagnetic and antiferromagnetic interactions. In the presence of carriers, T_C is given by the expression [80, 85]

$$T_C = \lfloor N_O X\text{eff} \cdot S(S+1)\beta^2 A_F P_S(T_C)/12k_B \rfloor - T_{AF}$$

where $N_O X$eff is the effective spin concentration, S is the localized spin state, β is the p–d exchange integral, A_F is the Fermi liquid parameter, P_S is the total density of states, k_B is Boltzmann's constant, and T_{AF} describes the contribution of antiferromagnetic interactions. Numerous refinements of this approach have appeared recently, taking into account the effects of positional disorder [86, 87], indirect exchange interactions [88], spatial inhomogeneities, and free-carrier spin polarization [89, 90]. Figure 2 shows a compilation of the predicted T_C values, together with the classification of the materials (e.g., group IV semiconductors, etc.). In the subsequent period after appearance of the Dietl et al. [80] paper, remarkable progress has been made on the realization of materials with T_C values at or above room temperature.

The mean field model and its variants produces reliable estimates of T_C for materials such as (Ga,Mn)As and (In,Mn)As and predicts that (Ga,Mn)N will have a value above room temperature [80]. Examples of the predicted ferromagnetic transition temperatures for both (Ga,Mn)As and (Ga,Mn)N are shown in Figure 3 for four different variants of the mean field approach [91]. These are the standard mean field theory (T_C^{MF}), a version that accounts for the role of Coulomb interactions with holes in the valence band (exchange-enhanced, T_C^{X}), another version that accounts for correlations in Mn ion orientations (collective, T_C^{coll}), or an estimate based on where excited spin waves cancel out the total spin of the ground state (T_C^{est}) [91]. Note that the dependence of any of the calculated T_C values on hole density in the material is much steeper for (Ga,Mn)As than for (Ga,Mn)N. The range of predicted values for GaAs has a much higher distribution than for GaN. These data emphasize the point that the mean field theories produce fairly reliable predictions for (Ga,Mn)As but at this stage are not very accurate for (Ga,Mn)N.

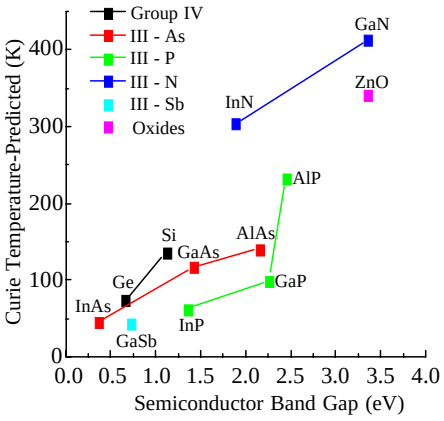


Figure 2. Predicted Curie temperatures as a function of bandgap (after [80]).

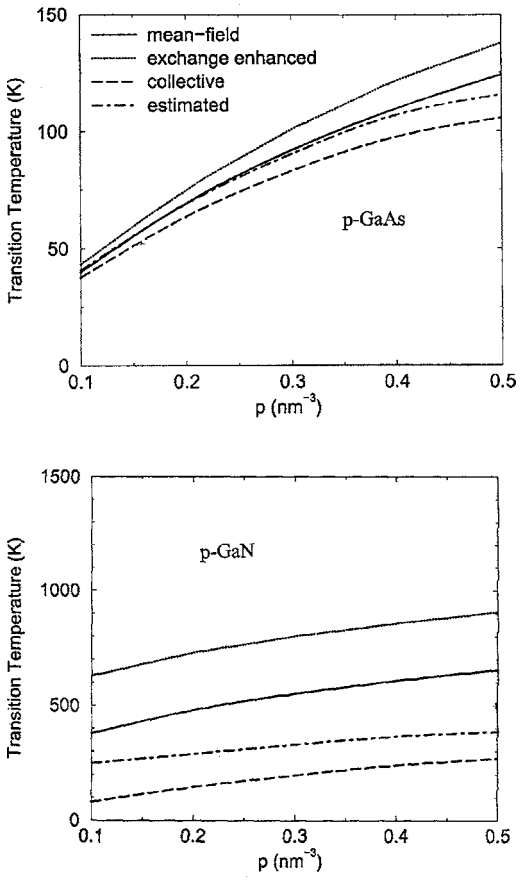


Figure 3. Predicted ferromagnetic transition temperatures in (Ga,Mn)As (left) or (Ga,Mn)N containing 5 at.% Mn, as a function of hole density. The four different curves in each graph represent results obtained from different variants of mean field theory (after [91]).

A second point largely overlooked in the theoretical work to date is that fact that the assumed hole densities may not be realistic. While GaAs can be readily doped with shallow acceptors such as C to produce hole densities of around 10^{21} cm^{-3} [92] and the Mn acceptors also contribute holes, the p-doping levels in GaN are limited to much lower values under normal conditions. For example, the ionization level (E_a) of the most common acceptor dopant in GaN, namely Mg, is relatively deep in the gap ($E_V + 0.17$ eV). Since the number of holes (P) is determined by the fraction of acceptors that are actually ionized at a given temperature T through a Boltzmann factor

$$P \propto \exp(-E_a/kT)$$

then for Mg at room temperature only a few percent of acceptors are ionized. While the Mg acceptor concentration in GaN can exceed 10^{19} cm^{-3}, a typical hole concentration at 25 °C is $P \sim 3 \times 10^{17}$ cm^{-3}. Initial reports of the energy level of Mn in GaN show it is very deep in the gap, $E_v + 1.4$ eV [93] and thus would be an ineffective dopant under most conditions. Some strategies for enhancing the hole concentration do exist, such as co-doping both acceptors and donors to reduce self-compensation effects [94] or the use of selectively doped AlGaN/GaN superlattices in which there is transfer of free holes from Mg acceptors in the AlGaN barriers to the GaN quantum wells [95]. These methods appear capable under optimum conditions of increasing the hole density in GaN to $>10^{18}$ cm^{-3} at 25 °C.

A further issue that needs additional exploration in the theories is the role of electrons, rather than holes, in stabilizing the ferromagnetism in DMS materials. All of the reports of ferromagnetism in (Ga,Mn)N, for example, occur for material that is actually *n*-type. Since the material has to be grown at relatively low temperatures to avoid Mn precipitation and therefore only MBE can be used, there is always the possibility of unintentional *n*-type doping from nitrogen vacancies, residual lattice defects, or impurities such as oxygen. Therefore stoichiometry effects, crystal defects, or unintentional impurities may control the final conductivity, rather than Mn or the intentionally introduced acceptor dopants. Once again, this is much less of an issue in materials such as GaAs, whose low temperature growth is relatively well understood and controlled.

While most of the theoretical work for DMS materials has focused on the use of Mn as the magnetic dopant, there has been some progress on identifying other transition metal atoms that may be effective. Figure 4 shows the predicted stability of ferromagnetic states in GaN doped with different 3*d* transition metal atoms [96]. The results are based on a local spin-density approximation which assumed that Ga atoms were randomly substituted with the magnetic atoms and did not take into account any additional carrier doping effects. In this study it was found that (Ga,V)N and (Ga,Cr)N showed stable ferromagnetism for all transition metal concentrations whereas Fe, Co, or Ni doping produced spin-glass ground states [96]. For the case of Mn, the ferromagnetic state was the lowest energy state for concentrations up to ~20% , whereas the spin-glass state became the most stable at higher Mn concentrations.

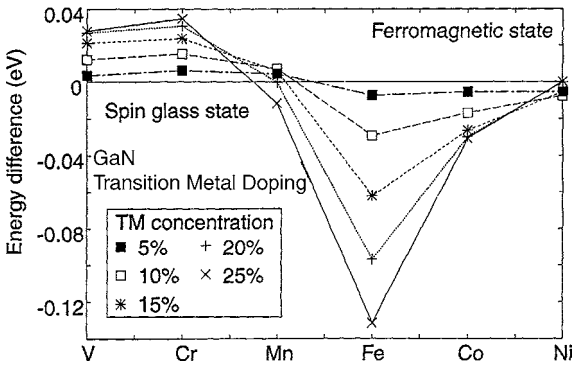


Figure 4. Predicted stability of the ferromagnetic states of different transition metal (TM) atoms in GaN as a function of transition metal concentration. The vertical axis represents the energy difference between the ferromagnetic and spin-glass states for each metal atom (after [96]).

4. EXPERIMENTAL RESULTS

4.1. (Ga,Mn)P

Ferromagnetism above room temperature in (Ga,Mn)P has been reported for two different methods of Mn incorporation, namely ion implantation [97] and doping during MBE growth [97, 98]. The implantation process is an efficient one for rapidly screening whether particular combinations of magnetic dopants and host semiconductors are promising in terms of ferromagnetic properties. We have used implantation to introduce ions such as Mn, Fe, and Ni into a variety of substrates, including GaN, SiC, and GaP

The temperature-dependent magnetization of a strongly *p*-type ($p \sim 10^{20}$), carbon-doped GaP sample implanted with ~6 at.% of Mn and then annealed at 700 °C is shown in Figure 5. The diamagnetic contribution was subtracted from the background. A Curie temperature (T_C) of ~270 K is indicated by the dashed vertical line, while the inset shows a ferromagnetic Curie temperature of 236 K.

Examples of hysteresis loops from MBE grown samples are shown in Figure 6. The hysteresis could be detected to 330 K. No secondary phases (such as MnGa or MnP) or clusters were determined by transmission electron microscopy, X-ray diffraction, or selected-area diffraction pattern analysis.

While mean-field theories predict relatively low Curie temperatures (<110 K) for (Ga,Mn)P [97, 98], recent experiments show ferromagnetism above 300 K [97, 98]. In other respects, the magnetic behavior of the (Ga,Mn)P was consistent with mean-field predictions. For example, the magnetization versus temperature plots showed a more classical concave shape than observed with many DMS materials.

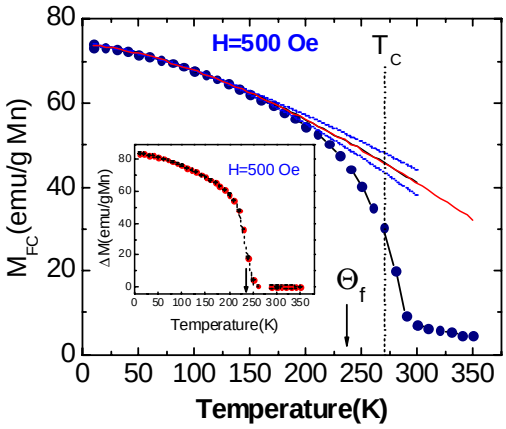


Figure 5. Field-cooled magnetization of (Ga,Mn)P as a function of temperature. The solid line shows a Bloch law dependence, while the dashed lines are 95% confidence bands. The vertical dashed line at $T_C = 270$ K is the field-independent inflection point and the vertical arrows in the main panel and inset mask to ferromagnetic Curie temperature Θ_f. The inset shows the temperature dependence of difference in magnetization between field-cooled and zero-field-cooled conditions.

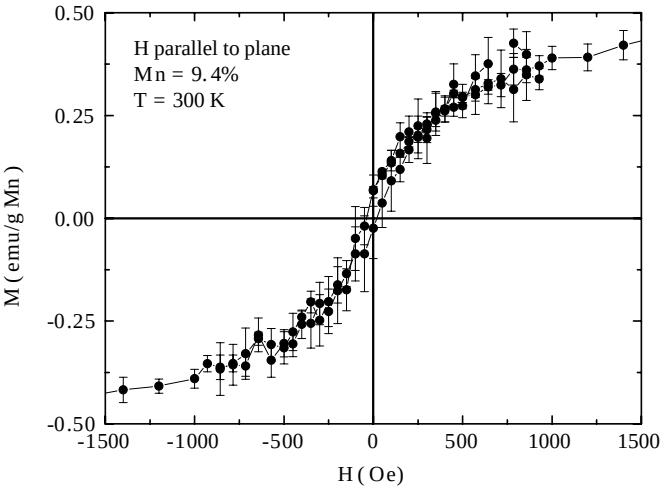


Figure 6. Magnetization versus field for MBE grown (Ga,Mn)P with 9.4 at.% Mn.

In addition, the Curie temperature was strongly influenced by the carrier density and type in the material, with highly *p*-type samples showing much higher values than *n*-type or undoped samples. Finally, the Curie temperature increased with Mn concentration up to ~6 at.% and decreased at higher concentrations [98]. No secondary phases or clusters could be detected by transmission electron microscopy, X-ray diffraction, or selected-area diffraction patterns. Similar results were achieved in samples in which the Mn was incorporated during MBE growth or directly implanted with Mn.

GaP is a particularly attractive host material for spintronic applications because it is almost lattice-matched to Si. One can therefore envision integration of (Ga,Mn)P spintronic magnetic sensors or data storage elements to form fast nonvolatile magnetic random access memories. Although it has an indirect bandgap, it can be made to luminescence through addition of isoelectric dopants such as nitrogen or else one could employ the direct bandgap ternary InGaP, which is lattice-matched to GaAs. The quaternary InGaAlP materials system is used for visible light-emitting diodes, laser diodes, heterojunction bipolar transistors, and high electron mobility transistors. An immediate application of the DMS counterparts to the component binary and ternary materials in this system would be to add spin functionality to all of these devices. A further advantage to the wide bandgap phosphides is that they exhibit room temperature ferromagnetism even for relatively high growth temperatures during MBE.

Obviously, the Mn can also be incorporated during MBE growth of the (Ga,Mn)P. The *p*-type doping level can be separately controlled by incorporating carbon from a CBr_4 source while P is obtained from thermal cracking of PH_3. A phase diagram for the epi growth of this materials system has been developed and this can be used to tailor the magnetic properties of the (Ga,Mn)P. For samples grown at 600 °C with 9.4 at.% Mn, hysteresis is still detectable at 300 K, with a coercive field of ~39 Oe.

4.2. (Ga,Mn)N

The initial work on this material involved either microcrystals synthesized by nitridization of pure metallic Ga in supercritical ammonia or bulk crystals grown in reactions of Ga/Mn alloys on GaN/Mn mixtures with ammonia at ~1200 °C [99]. These samples exhibit paramagnetic properties over a broad range of Mn concentrations, as did some of its early MBE-grown films. By contrast, Figure 7 shows room temperature ferromagnetism from more recent *n*-type (Ga,Mn)N samples.

The first reports of the magnetic properties of (Ga,Mn)N involved bulk microcrystallites grown at high temperatures (~1200 °C), but while percent levels of Mn were incorporated, the samples exhibited paramagnetic behavior [99]. By sharp contrast, in epitaxial GaN layers grown on sapphire substrates and then subjected to solid state diffusion of Mn at temperatures from 250 to 800 °C for various periods, clear signatures of room temperature ferromagnetism were observed [100, 101]. Figure 8 shows anomalous Hall effect data (left) at 323 K and the temperature dependence of sheet resistance at zero applied field for two different Mn-diffused samples and an undoped GaN control sample (right) [101]. The Curie temperature was found to be in the range 220–370 K, depending on the diffusion conditions. The use of ion implantation to introduce the Mn produced lower magnetic ordering temperatures [102].

In (Ga,Mn)N films grown by MBE at temperatures between 580 and 720 °C with Mn contents of 6–9 at.%, magnetization (M) versus magnetic field (H) curves showed clear hysteresis at 300 K, with coercivities of 52–85 Oe and residual magnetizations of 0.08–0.77 emu/g at this

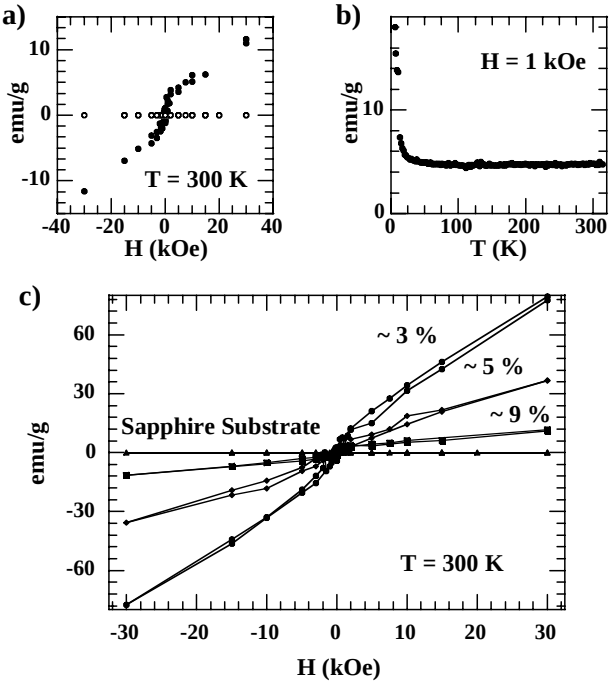


Figure 7. (a) B–H from MBE grown (Ga,Mn)N with 9.4 at.% Mn (closed circles) and from sapphire substrate (upper circles), (b) M–T of (Ga,Mn)N, (c) B–H from (Ga,Mn)N as a function of Mn concentration.

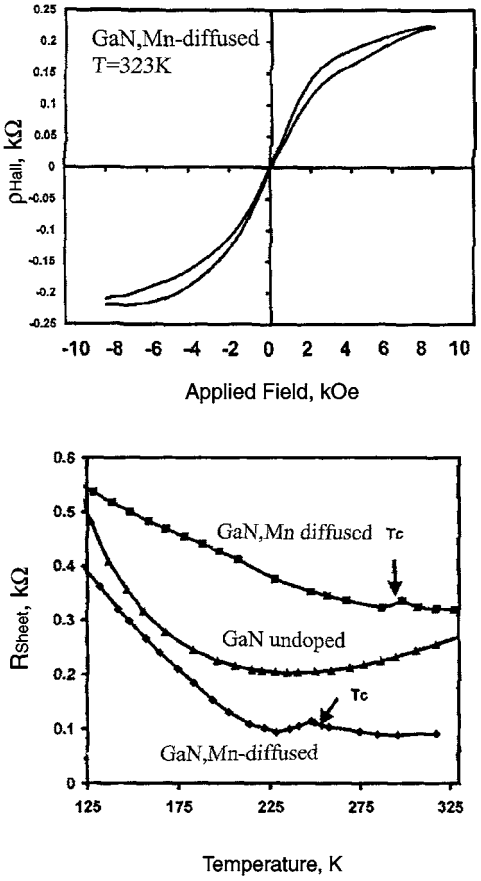


Figure 8. Temperature dependence of sheet resistance at zero magnetic field for Mn-diffused GaN and as-grown GaN (bottom) and room temperature anomalous Hall effect hysteresis curves for Mn-diffused GaN (top) (after [101]).

temperature [103]. Figure 9 shows the temperature dependence of the magnetization for a sample with 9 at.% Mn, yielding an estimated T_C of 940 K using a mean field approximation. Note that while the electrical properties of the samples were not measured, they were almost certainly *n*-type.

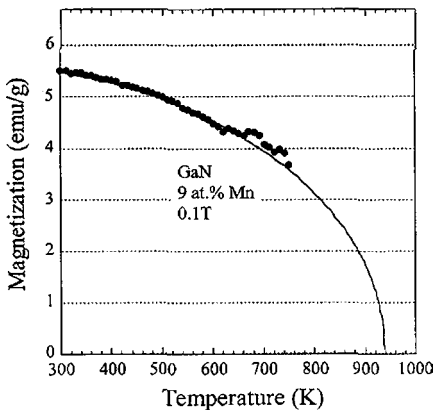


Figure 9. Magnetization versus temperature for (Ga,Mn)N sample grown by MBE with ~9 at.% Mn. The extrapolation of the curve is based on a mean field approximation (after [103]).

As we discussed above, it is difficult to obtain high Curie temperatures in *n*-type DMS materials according to the mean-field theories and this is something that needs to be addressed in future refinements of these theories. Room temperature ferromagnetism in *n*-type (Ga,Mn)N grown by MBE has also been reported by Thaler et al. [104]. In that case, strenuous efforts were made to exclude any possible contribution from the sample holder in the SQUID magnetometer or other spurious effects. It is also worthwhile to point out that for the studies of (Ga,Mn)N showing ferromagnetic ordering by magnetization measurements, a number of materials characterization techniques did not show the presence of any second ferromagnetic phases within detectable limits. In addition, the values of the measured coercivities are relatively small. If indeed there were undetectable amounts of nanosized clusters, due to geometrical effects, the expected fields at which these clusters would switch magnetically would be expected to be much larger than what has been observed. EXAFS measurements performed on (Ga,Mn)N samples grown by MBE on sapphire at temperatures of 400–650 °C with Mn concentrations of $\sim 7 \times 10^{20}$ cm^{-3} (i.e., slightly over 2 at.%) are shown in Figure 10 [105]. The similarity of the experimental data with simulated curves for a sample containing this concentration of Mn substituted for Ga on substitutional lattice positions indicates that Mn is in fact soluble at these densities. In the samples grown at 650 °C, $\leq$1 at.% of the total amount of Mn was found to be present as Mn clusters. However, at lower growth temperatures (400 °C), the amount of Mn that could be present as clusters increased up to ~36 at.% of the total Mn incorporated. The ionic state of the substitutional Mn was found to be primarily Mn(2), so that these impurities act as acceptors when substituting for the Ga with valence three. However, when the electrical properties of these samples were measured, they were found to be resistive [105]. This result emphasizes how much more needs to

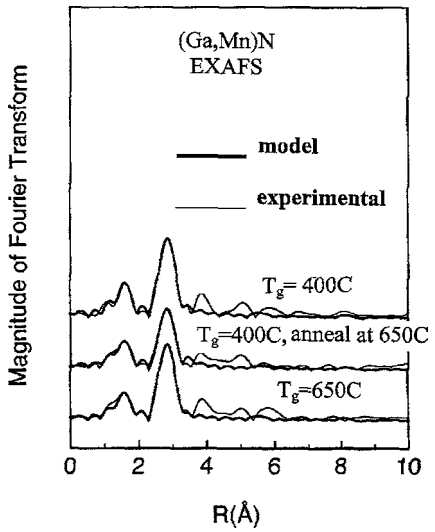


Figure 10. Fourier transforms of the Mn *K*-edge EXAFS data from Mn-doped (GaMn)N. The fine lines are the experimental data in these curves, while the course lines are the calculated curves assuming 2 at.% of Mn in the GaN (after [105]).

be understood concerning the effects of compensation and unintentional doping of (Ga,Mn)N, since the EXAFS data indicated the samples should have shown very high *p*-type conductivity due to incorporation of Mn acceptors.

Other reports have also recently appeared on the magnetic properties of GaN doped with other transition metal impurities. For initially *p*-type samples directly implanted with either Fe or Ni, ferromagnetism was observed at temperatures of ~200 [106] and 50 K [107], respectively. (Ga,Fe)N films grown by MBE showed Curie temperatures of ≤100 K, with EXAFS data showing that the majority of the Fe was substitutional on Ga sites [108]. (Ga,Cr)N layers grown in a similar fashion at 700 °C on sapphire substrates showed single-phase behavior, clear hysteresis and saturation of magnetization at 300 K, and a Curie temperature exceeding 400 K [109].

Epi growth of (Ga,Mn)N has produced a range of growth conditions producing single-phase material and the resulting magnetic properties [100–107]. In general, no second phases are found for Mn levels below ~10% for growth temperatures of ~750 °C. The (Ga,Mn)N retains *n*-type conductivity under these conditions.

In accordance with most of the theoretical predictions, magnetotransport data showed the anomalous Hall effect, negative magnetoresistance, and magnetic resistance at temperatures that were dependent on the Mn concentration. For example, in films with very low (<1%) or very high (~9%) Mn concentrations, the Curie temperatures were between 10 and 25 K. An example is shown in Figure 7 for an *n*-type (Ga,Mn)N sample with Mn ~7%. The sheet resistance shows negative magnetoresistance below 150 K, with the anomalous Hall coefficient disappearing below 25 K. When the Mn concentration was decreased to 3 at.%, the (Ga,Mn)N showed the highest degree of ordering per Mn atom [104]. Figure 7a shows hysteresis present at 300 K, while the magnetization as a function of temperature is shown in Figure 7b. Data from samples with different Mn concentrations are shown in Figure 7c and indicate ferromagnetic coupling, leading to a lower moment per Mn. Data from field-cooled and zero-field—cooled conditions were further suggestive of room temperature magnetization [104]. The significance of these results is that there are many advantages from a device viewpoint to having *n*-type ferromagnetic semiconductors.

The local structure and effective chemical valency of Mn in MBE-grown (Ga,Mn)N samples has been investigated by EXAFS [105]. It was concluded that most of the Mn was incorporated substitutionally on the Ga sublattice with effective valency close to +2 for samples with ~2 at.% Mn [105]. There was also evidence that a fraction (from 1% to 36%, depending on growth condition) of the total Mn concentration could be present as small Mn clusters [105].

4.3. Chalcopyrite Materials

The chalcopyrite semiconductors are of interest for a number of applications. For example, $ZnGeP_2$ exhibits unusual nonlinear optical properties and can be used in optical oscillators and frequency converters. $ZnSnAs_2$ shows promise for far-IR generation and frequency converters. The wide bandgap chalcopyrites $ZnGeN_2$ and $ZnSiN_2$ have lattice parameters close to GaN and SiC, respectively, and the achievement of ferromagnetism in these materials would make it possible for direct integration of magnetic sensors and switches with blue/green/ultraviolet (UV) lasers and light-emitting diodes, UV solar-blind detectors, and microwave power electronic devices fabricated in the GaN and SiC. The bandgap of $ZnGe_xSi_{1-x}N_2$ varies linearly with composition from 3.2 ($x = 1$) to 4.46 eV ($x = 0$).

Numerous reports of room temperature ferromagnetism in Mn-doped chalcopyrites have appeared. A compilation of these results and those from transition metal doped GaN and GaP are shown in Table 1. The $ZnSnAs_2$ is somewhat of an anomaly due to its small bandgap, but little theory is available at this point on the chalcopyrites and their expected magnetic properties as a function of bandgap, doping, or Mn concentration. In the only case in which electrical properties were reported, the $ZnGeSiN_2$:Mn was *n*-type [110]. There is also no information available on the energy level of Mn in the bandgap.

Table 1. Compilation of semiconductors exhibiting room temperature ferromagnetism.

Material	Bandgap (eV)	Synthesis	T_C(K)	Ref.
$Cd_{1-x}Mn_xGeP_2$	1.72	Solid-phase reaction of evap. Mn	>300	[74]
(Ga,Mn)N	3.4	Mn incorporated by diff^n	228–370	[100, 101]
(Ga,Mn)N	3.4	Mn incorporated during MBE; n-type	>300	104
(Ga,Mn)N	3.4	Mn incorporated during MBE	940	[103]
(Ga,Cr)N	3.4	Cr incorporated during MBE	>400	[109]
(Ga,Mn)P:C	2.2	Mn incorporated by implant or MBE; p ~10^{20} cm^{-3}	>330	[97, 98]
$(Zn_{1-x}Mn_x)GeP_2$	1.83–2.8	Sealed ampule growth; insulating; 5.6% Mn	312	[74]
$(ZnMn)GeP_2$	<2.8	Mn incorporated by diff.	350K	[75]
$ZnSnAs_2$	0.65	Bridgman bulk growth	329K	[76]
$ZnSiGeN_2$	3.52	Mn-implanted epi	~300	[106]

Note: Extrapolated from measurements up to ~750 K.

5. POTENTIAL DEVICE APPLICATIONS

Previous articles have discussed some spintronic device concepts such as spin junction diodes and solar cells [9], optical isolators, and electrically controlled ferromagnets [10]. The realization of light-emitting diodes with a degree of polarized output has been used to measure spin injection efficiency in heterostructures [111–113]. Such structures can reveal much about spin transport through heterointerfaces after realistic device processing schemes involving etching, annealing, and metallization. The spin transfer in such situations has proven surprisingly robust [114]. It is obviously desirable that spintronic devices are operable at or above room temperature. As an initial demonstration that (Ga,Mn)N layers can be used as the *n*-type injection layer in GaN/InGaN blue light-emitting diodes (LEDs), Figure 11 shows the LED during operation (top) and the spectral output (left). It is necessary to next establish the extent of any degree of polarization of the light emission, which might be difficult to observe in GaN/InGaN LEDs, since it has been shown that the free exciton components in the electroluminescence spectrum contribute mostly to the observed circular polarization of the emitted light [115]. While the expected advantages of spin-based devices include nonvolatility, higher integration densities, lower power operation, and higher switching speeds, there are many factors still to consider in whether any of these can be realized. These factors include whether the signal sizes due to spin effects are large enough at room temperature to justify the extra development work needed to make spintronic devices and whether the expected added functionality possible will materialize.

In addition to active and/or optical devices, wide bandgap DMS materials may also be used as passive devices. LeClair et al. [116] have recently shown an artificial half-metallic structure by using a polycrystalline sputtered ferromagnetic semiconductor (EuS) as a tunneling barrier. This barrier can function as an effective spin filter, since a tunneling electron encounters a differing barrier height depending on its spin below T_c of the barrier material (for EuS, T_C ~16.8 K). At low temperatures, the spin-filtering efficiencies were found to be ~90%. Room temperature DMS materials for these spin-filtering effects could be used to increase magnetoresistance changes in current magnetic tunnel junctions and metallic spin-valve structures.

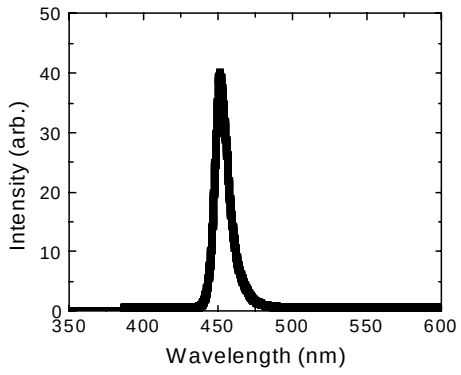


Figure 11. Spectral output from GaMnN/InGaN light-emitting diode.

6. ISSUES TO BE RESOLVED

As described earlier, there are a number of existing models for the observed ferromagnetism in semiconductors. The near-field models consider the ferromagnetism to be mediated by delocalized or weakly localized holes in the *p*-type materials. The magnetic Mn ion provides a localized spin and acts as an acceptor in most III–V semiconductors so that it can also provide holes. In these models, the T_C is proportional to the density of Mn ions and the hole density. Many aspects of the experimental data can be explained by the basic mean field model. However, ferromagnetism has been observed in samples that have very low hole concentrations, in insulating material, and more recently in *n*-type material. Models in these regimes are starting to appear [117].

An alternative approach using local density functional calculations suggests that the magnetic impurities may form small nanosize clusters that produce the observed ferromagnetism [82]. These clusters would be difficult to detect by most characterization techniques. Clearly there is a need to more fully characterize the materials showing room temperature ferromagnetism and correlate these results to establish on a case-by-case basis which is the operative mechanism and also to refine the theories based on experimental input. More work is also needed to establish the energy levels of the Mn, whether there are more effective magnetic dopant atoms, and how the magnetic properties are influenced by carrier density and type. Even basic measurements such as how the bandgap changes with Mn concentration in GaN and GaP have not been performed. The control of spin injection and manipulation of spin transport by external means such as voltage from a gate contact or magnetic fields from adjacent current lines or ferromagnetic contacts is at the heart of whether spintronics can be exploited in device structures, and these areas are still in their infancy. A concerted effort on the physics and materials science of the new dilute magnetic semiconductors is underway in many groups around the world, but fresh insights, theories, and characterization methods would greatly accelerate the process.

GLOSSARY

Dilute magnetic semiconductor A semiconductor in which some of the lattice atoms are replaced by magnetic ions, leading of ferromagnetism.

GaMnN Gallium nitride doped with a high concentration (a few percent) of Mn, to induce ferromagnetism.

Spin-fets A field effect transistor in which the spin of the electron is used to control the current flow.

Spin-led A light-emitting diode, in which at least one side of the junction is doped with Mn to produce polarized carriers, leading to polarized light emission.

Spintronics Spin transfer electronics.

ACKNOWLEDGMENTS

The work at UF was partially supported by NSF-DMR 0101438, while the work at SNU was partially supported by KOSEF and Samsung Electronics Endowment through

CSCMR and by the Seoul National University Research Foundation. The authors are very grateful to their collaborators M. E. Overberg, G. T. Thaler, F. Ren, Jihyun Kim, A. F. Hebard, N. A. Theodoropoulou, R. G. Wilson, J. M. Zavada, D. P. Norton, S. N. G. Chu, J. S. Lee, and Z. G. Khim.

REFERENCES

1. See for example http://www.almaden.ibm.com/sst/.
2. H. Ohno, D. Chiba, F. Matsukura, T. Omiya, E. Abe, T. Dietl, Y. Ohno, and K. Ohtani, *Nature* 408, 944 (2000).
3. A. Oiwa, Y. Mitsumori, R. Moriya, T. Slupinski, and H. Munekata, *Phys. Rev. Lett.* 88, 137202 (2002).
4. Y. D. Park, A. T. Hanbicki, S. C. Erwin, C. S. Hellberg, J. M. Sullivan, J. E. Mattson, A. Wilson, G. Spanos, and B. T. Jonker, *Science* 295, 651 (2002).
5. H. Ohno, *J. Vac. Sci. Technol. B* 18, 2039 (2000).
6. S. A. Wolf, D. D. Awschalom, R. A. Buhrman, J. M. Daughton, S. von Molnar, M. L. Roukes, A. Y. Chtchelkanova, and D. M. Treger, *Science* 294, 1488 (2001).
7. S. Von Molnar et al., "World Technology (WTEC) Study on Spin Electronics: Highlights of Recent US Research and Development Activities," http://www.wtec.org/spin_US_summary.pdf (2001).
8. S. A. Chambers, *Mater. Today*, pp. 34–39, April 2002.
9. S. Das Sarma, *Am. Scientist* 89, 516 (2001).
10. H. Ohno, F. Matsukura, and Y. Ohno, *JSAP Int*. 5, 4 (2002).
11. D. D. Awschalom and J. M. Kikkawa, *Science* 287, 473 (2000).
12. C. Gould, G. Schmidt, G. Richler, R. Fiederling, P. Grabs, and L. W. Molenkamp, *Appl. Surf. Sci.*, 60, 172 (2002).
13. P. R. Hammar, B. R. Bennett, M. J. Yang, and Mark Johnson, *Phys. Rev. Lett.* 83, 203 (1999).
14. F. G. Monzon, H. X. Tang, and M. L. Roukes, *Phys. Rev. Lett.* 84, 5022 (2000).
15. G. A. Prinz, Magnetic metal films on semiconductor substrates, in "Ultrathin Magnetic Structures II" (B. Heinrich and J. A. C. Bland, Eds.). Springer-Verlag, New York, 1994.
16. G. Schmidt, D. Ferrand, L. W. Molenkamp, A. T. Filip, and B. J. van Wees, *Phys. Rev. B* 62, R4793 (2000).
17. E. I. Rashba, *Phys. Rev. B* 62, R16267 (2000).
18. H. J. Zhu, M. Ramsteiner, H. Kostial, M. Wassermeier, H.-P. Schönherr, and K. H. Ploog, *Phys. Rev. Lett.* 87, 016601 (2001).
19. F. Matsukura, H. Ohno, A. Shen, and Y. Sugawara, *Phys. Rev. B* 57, R2037 (1998).
20. H. Ohno, A. Shen, F. Matsukura, A. Oiwa, A. Endo, S. Katsumoto, and Y. Iye, *Appl. Phys. Lett.* 69, 363 (1996).
21. R. Shioda, K. Ando, T. Hayashi, and M. Tanaka, *Phys. Rev. B* 58, 1100 (1998).
22. Y. Satoh, N. Inoue, Y. Nishikawa, and J. Yoshino, in "Proc. 3rd Symp. Physics and Applications of Spin-Related Phenomena in Semiconductors" (H. Ohno, J. Yoshino, and Y. Oka, Eds.), Sendai, Japan, 1997, p. 23.
23. T. Hayashi, M. Tanaka, T. Nishinaga, H. Shimoda, H. Tsuchiya, and Y. Otsuka, *J. Cryst. Growth* 175, 1063 (1997).
24. A. Van Esch, L. Van Bockstal, J. de Boeck, G. Verbanck, A. S. vas Steenbergen, R. J. Wellman, G. Grietens, R. Bogaerts, F. Herlach, and G. Borghs, *Phys. Rev. B* 56, 13103 (1997).
25. B. Beschoten, P. A. Crowell, I. Malajovich, D. D. Awschalom, F. Matsukura, A. Shen, and H. Ohno, *Phys. Rev. Lett.* 83, 3073 (1999).
26. M. Tanaka, *J. Vac. Sci. Technol. B* 16, 2267 (1998).
27. Y. Nagai, T. Kunimoto, K. Nagasaka, H. Nojiri, M. Motokawa, F. Matsukura, T. Dietl, and H. Ohno, *Japan. J. Appl. Phys.* 40, 6231 (2001).
28. J. Sadowski, R. Mathieu, P. Svedlindh, J. Z. Domagala, J. Bak-Misiuk, J. Swiatek, M. Karlsteen, J. Kanski, L. Ilver, H. Asklund, and V. Sodervall, *Appl. Phys. Lett.* 78, 3271 (2001).
29. A. Shen, F. Matsukura, S. P. Guo, Y. Sugawara, H. Ohno, M. Tani, A. Abe, and H. C. Liu, *J. Cryst. Growth* 201/202, 379 (1999).
30. H. Shimizu, T. Hayashi, T. Nishinaga, and M. Tanaka, *Appl. Phys. Lett.* 74, 398 (1999).
31. B. Grandidier, J. P. Hys, C. Delerue, D. Stievenard, Y. Higo, and M. Tanaka, *Appl. Phys. Lett.* 77, 4001 (2000).
32. R. K. Kawakami, E. Johnson-Halperin, L. F. Chen, M. Hanson, N. Guebels, J. S. Speck, A. C. Gossard, and D. D. Awschalom, *Appl. Phys. Lett.* 77, 2379 (2000).
33. K. Ando, T. Hayashi, M. Tanaka, and A. Twardowski, *J. Appl. Phys.* 83, 65481 (1998).
34. D. Chiba, N. Akiba, F. Matsukura, Y. Ohno, and H. Ohno, *Appl. Phys. Lett.* 77, 1873 (2000).
35. H. Ohno, F. Matsukura, T. Owiya, and N. Akiba, *J. Appl. Phys.* 85, 4277 (1999).
36. T. Hayashi, M. Tanaka, T. Nishinaga, and H. Shimada, *J. Appl. Phys.* 81, 4865 (1997).
37. T. Hayashi, M. Tanaka, K. Seto, T. Nishinaga, and K. Ando, *Appl. Phys. Lett.* 71, 1825 (1997).
38. A. Twardowski, *Mater. Sci. Eng. B* 63, 96 (1999).
39. T. Hayashi, M. Tanaka, and A. Asamitsu, *J. Appl. Phys.* 87, 4673 (2000).
40. N. Akiba, D. Chiba, K. Natata, F. Matsukura, Y. Ohno, and H. Ohno, *J. Appl. Phys.* 87, 6436 (2000).
41. S. J. Potashnik, K. C. Ku, S. H. Chun, J. J. Berry, N. Samarth, and P. Schiffer, *Appl. Phys. Lett.* 79, 1495 (2001).
42. G. M. Schott, W. Faschinger, and L. W. Molenkamp, *Appl. Phys. Lett.* 79, 1807 (2001).
43. H. Munekata, H. Ohno, S. von Molnar, A. Segmuller, L. L. Chang, and L. Esaki, *Phys. Rev. Lett.* 63, 1849 (1989).
44. H. Akai, *Phys. Rev. Lett.* 81, 3002 (1998).
45. H. Ohno, H. Munekata, T. Penney, S. von Molnar, and L. L. Chang, *Phys. Rev. Lett.* 68, 2864 (1992).
46. H. Munekata, A. Zaslevsky, P. Fumagalli, and R. J. Gambino, *Appl. Phys. Lett.* 63, 2929 (1993).
47. S. Koshihara, A. Oiwa, M. Hirasawa, S. Katsumoto, Y. Iye, C. Urano, H. Takagi, and H. Munekata, *Phys. Rev. Lett.* 78, 4617 (1997).
48. Y. L. Soo, S. W. Huang, Z. H. Ming, Y. H. Kao, and H. Munekata, *Phys. Rev. B* 53, 4905 (1996).
49. A. Oiwa, T. Slupinski, and H. Munekata, *Appl. Phys. Lett.* 78, 518 (2001).
50. Y. Nishikawa, A. Tackeuchi, M. Yamaguchi, S. Muto and O. Wada, *IEEE J. Sel. Topics Quantum Electron*. 2, 661 (1996).
51. H. Munekata, presented at ICCG-13, August 2001.
52. I. Malajovich, J. M. Kikkawa, D. D. Awschalom, J. J. Berry, and N. Samarth, *Phys. Rev. Lett.* 84, 1015 (2000).
53. P. R. Hammar, B. R. Bennet, M. Y. Yang, and M. Johnson, *J. Appl. Phys.* 87, 4665 (2000).
54. A. Hirohata, Y. B. Xu, C. M. Guetler, and J. A. C. Bland, *J. Appl. Phys.* 87, 4670 (2000).
55. M. Johnson, *J. Vac. Sci. Technol. A* 16, 1806 (1998).
56. S. Cardelis, C. G. Smith, C. H. W. Barnes, E. H. Linfield, and J. Ritchie, *Phys. Rev. B* 60, 7764 (1999).
57. R. Fiederling, M. Kein, G. Reresher, W. Ossan, G. Schmidt, A. Wang, and L. W. Molenkamp, *Nature* 402, 787 (1999).
58. G. Borghs and J. De Boeck, *Mater. Sci. Eng. B* 84, 75 (2001).
59. Y. Ohno, D. K. Young, B. Beschoten, F. Matsukura, H. Ohno, and D. D. Awschalom, *Nature* 402, 790 (1999).
60. B. T. Jonker, Y. D. Park, B. R. Bennett, H. D. Cheong, G. Kiseoglou, and A. Petrou, *Phys. Rev. B* 62, 8180 (2000).
61. Y. D. Park, B. T. Jonker, B. R. Bennett, G. Itskos, M. Furis, G. Kioseoglou, and A. Petrou, *Appl. Phys. Lett.* 77, 3989 (2000).

62. G. Schmidt, D. Ferrand, L. W. Molenkamp, A. T. Filip, and B. J. van Wees, *Phys. Rev. B* 62, R4790 (2000).
63. Y. Q. Jin, R. C. Shi, and S. J. Chou, *IEEE Trans. Magn.* 32, 4707 (1996).
64. C. M. Hu, J. Nitta, A. Jensen, J. B. Hansen, and H. Takayanagai, *Phys. Rev. B* 63, 125333 (2001).
65. F. G. Monzon, H. X. Tang, and M. L. Roukes, *Phys. Rev. Lett.* 84, 5022 (2000).
66. S. Gardelis, C. G. Smith, C. H. W. Barnes, F. Matsukura, and H. Ohno, *Japan. J. Appl. Phys.* 40, L1274 (2001).
67. H. Breve, S. Nemeth, Z. Liu, J. De Boeck, and G. Borghs, *J. Magn. Magn. Mater*. 226–230, 933 (2001).
68. H. Munekata, H. Ohno, S. von Molnar, A. Segmuller, L. L. Chang, and L. Esaki, *Phys. Rev. Lett.* 63, 1849 (1989).
69. H. J. Zhu, M. Ramsteiner, H. Kostial, M. Wassermeier, H. P. Schononherr, and K. H. Ploog, *Phys. Rev. Lett.* 87, 016601 (2001).
70. M. Kohda, Y. Ohno, K. Takamura, F. Matsukura, and H. Ohno, *Japan. J. Appl. Phys.* 40, L1274 (2001).
71. J. A. Katine, F. J. Albert, R. A. Buhrman, E. D. Myers, and D. C. Ralph, *Phys. Rev. Lett.* 84, 319 (2000).
72. J. Nitta, T. Ahazaki, H. Takayanngi, and T. Enoki, *Phys. Rev. Lett.* 78, 1335 (1997).
73. S. Cho, S. Choi, G. B. Cha, S. C. Hong, Y. Kim, Y.-J. Zhao, A. J. Freeman, J. B. Ketterson, B. J. Kim, Y. C. Kim, and B. C. Choi, *Phys. Rev. Lett.* 88, 257203-1 (2002).
74. G. A. Medvedkin, T. Ishibashi, T. Nishi, and K. Hiyata, *Japan. J. Appl. Phys.* 39, L949 (2000).
75. G. A. Medvedkin, K. Hirose, T. Ishibashi, T. Nishi, V. G. Voevodin, and K. Sato, K., *J. Cryst. Growth* 236, 609 (2002).
76. S. Choi, G. B. Cha, S. C. Hong, S. Cho, Y. Kim, J. B. Ketterson, S.-Y. Jeong, and G. C. Yi, *Solid State Commun.* 122, 165 (2002).
77. K. Ueda, H. Tahata, and T. Kawai, *Appl. Phys. Lett.* 79, 988 (2001).
78. Y. Matsumoto, M. Murakami, T. Shono, H. Hasegawa, T. Fukumura, M. Kawasaki, P. Ahmet, T. Chikyow, S. Koshikara, and H. Koinuma, *Science* 291, 854 (2001).
79. F. Holtzberg, S. von Molnar, and J. M. D. Coey, "Handbook on Semiconductors" (T. Moss, Ed.). North-Holland, Amsterdam, 1980.
80. T. Dietl, H. Ohno, F. Matsukura, J. Cibert, and D. Ferrand, *Science* 287, 1019 (2000).
81. C. Zener, *Phys. Rev. B* 81, 440 (1951).
82. M. Van Schilfgaarde and O. N. Myrasov, *Phys. Rev. B* 63, 233205 (2001).
83. T. Dietl, H. Ohno, and F. Matsukura, *Phys. Rev. B* 63, 195205 (2001).
84. T. Dietl, *J. Appl. Phys.* 89, 7437 (2001).
85. T. Jungwirth, W. A. Atkinson, B. Lee, and A. H. MacDonald, *Phys. Rev. B* 59, 9818 (1999).
86. M. Berciu and R. N. Bhatt, *Phys. Rev. Lett.* 87, 108203 (2001).
87. R. N. Bhatt, M. Berciu, M. D. Kennett, and X. Wan, *J. Superconductivity* 15, 71 (2002).
88. V. I. Litvinov and V. A. Dugaev, *Phys. Rev. Lett.* 86, 5593 (2001).
89. J. Konig, H. H. Lin, and A. H. MacDonald, *Phys. Rev. Lett.* 84, 5628 (2001).
90. J. Schliemann, J. Konig, and A. H. MacDonald, *Phys. Rev. B* 64, 165201 (2001).
91. T. Jungwirth, J. Konig, J. Sinova, J. Kucera, and A. H. MacDonald, *Phys. Rev. B* 65, 10082 (2002).
92. C. R. Abernathy, *Mater. Sci. Rep. R* 16, 203 (1995).
93. R. Y. Korortiev, J. M. Gregie, and B. W. Wessels, *Appl. Phys. Lett.* 80, 1731 (2002).
94. H. Katayama-Yoshida, R. Kato, and T. Yamamoto, *J. Cryst. Growth* 231, 438 (2001).
95. I. D. Goepfert, E. F. Schubert, A. Osinsky, P. E. Norris, and N. N. Faleev, *J. Appl. Phys.* 88, 2030 (2000).
96. K. Sato and H. Katayama-Yoshida, *Japan. J. Appl. Phys.* 40, L485 (2001).
97. N. Theodoropoulou, A. F. Hebard, M. E. Overberg, C. R. Abernathy, S. J. Pearton, S. N. G. Chu, and R. G. Wilson, *Phys. Rev. Lett.*, 89, 107203 (2002).
98. M. E. Overberg, B. P. Gila, G. T. Thaler, C. R. Abernathy, S. J. Pearton, N. Theodoropoulou, K. T. McCarthy, S. B. Arnason, A. F. Hebard, S. N. G. Chu, R. G. Wilson, J. M. Zavada, and Y. D. Park, *J. Vac. Sci. Technol. B* 20, 969 (2002).
99. M. Zajac, J. Gosk, M. Kaminska, A. Twardowski, T. Szyszko, and S. Podliasko, *Appl. Phys. Lett.* 79, 2432 (2001).
100. M. L. Reed, M. K. Ritums, H. H. Stadelmaier, M. J. Reed, C. A. Parker, S. M. Bedair, and N. A. El-Masry, *Mater. Lett.* 51, 500 (2001).
101. M. L. Reed, N. A. El-Masry, H. Stadelmaier, M. E. Ritums, N. J. Reed, C. A. Parker, J. C. Roberts, and S. M. Bedair, *Appl. Phys. Lett.* 79, 3473 (2001).
102. N. Theodoropoulou, A. F. Hebard, M. E. Overberg, C. R. Abernathy, S. J. Pearton, S. N. G. Chu, and R. G. Wilson, *Appl. Phys. Lett.* 78, 3475 (2001).
103. S. Sonoda, S. Shimizu, T. Sasaki, Y. Yamamoto, and H. Hori, *J. Cryst. Growth* 237–239, 1358 (2002).
104. G. T. Thaler, M. E. Overberg, B. Gila, R. Frazier, C. R. Abernathy, S. J. Pearton, J. S. Lee, S. Y. Lee, Y. D. Park, Z. G. Khim, J. Kim, and F. Ren, *Appl. Phys. Lett.* 80 3964 (2002).
105. Y. L. Soo, G. Kioseouglou, S. Kim, S. Huang, Y. H. Kaa, S. Kubarawa, S. Owa, T. Kondo, and H. Munekata, *Appl. Phys. Lett.* 79, 3926 (2001).
106. N. A. Theodoropoulou, A. F. Hebard, S. N. G. Chu, M. E. Overberg, C. R. Abernathy, S. J. Pearton, R. G. Wilson, and J. M. Zavada, *Appl. Phys. Lett.* 79, 3452 (2001).
107. S. J. Pearton, M. E. Overberg, G. Thaler, C. R. Abernathy, N. Theodoropoulou, A. F. Hebard, S. N. G. Chu, R. G. Wilson, J. M. Zavada, A. Y. Polyakov, A. Osinsky, and Y. D. Park, *J. Vac. Sci. Technol. A* 20, 583 (2002).
108. H. Akinaga, S. Nemeth, J. De Boeck, L. Nistor, H. Bender, G. Borghs, H. Ofuchi, and M. Oshima, *Appl. Phys. Lett.* 77, 4377 (2000).
109. M. Hashimoto, Y. Z. Zhou, M. Kanamura, and H. Asahi, *Solid-State Commun*. 122, 37 (2002).
110. S. J. Pearton, M. E. Overberg, C. R. Abernathy, N. A. Theodoropoulou, A. F. Hebard, S. N. G. Chu, A. Osinsky, V. Zuflyigin, L. D. Zhu, A. Y. Polyakov, and R. G. Wilson, *J. Appl. Phys.*, 92, 2047 (2002).
111. R. Fiederling, M. Kein, G. Resescher, W. Ossau, G. Schmidt, W. Wang, and L.W. Molenkamp, *Nature* 402, 787 (1999).
112. Y. Ohno, D. K. Young, B. Beschoten, F. Matsukura, H. Ohno, and D. D. Awschalom, *Nature* 402, 790 (1999).
113. B. T. Jonker, Y. D. Park, B. R. Bennet, H. D. Cheong, G. Kioseoglou, and A. Petrou, *Phys. Rev. B* 62, 8180 (2000).
114. Y. D. Park, B. T. Jonker, B. R. Bennet, G. Itzkos, M. Furis, G. Kioseoglou, and A. Petrou, *Appl. Phys. Lett.* 77, 3989 (2000).
115. B. T. Jonker, A. T. Hanbicki, Y. D. Park, G. Itskos, M. Furis, G. Kioseoglou, and A. Petrou, *Appl. Phys. Lett.* 79, 3098 (2001).
116. P. LeClair, J. K. Ha, H. J. M. Swagten, J. T. Kohlhepp, C. H. van de Vin, and W. J. M. de Jonge, *Appl. Phys. Lett.* 80, 625 (2002).
117. A. Kaminski and S. Das Sarma, *Phys. Rev. Lett.* 88, 247202-1 (2002).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanoscale Heat Transfer

G. Chen, D. Borca-Tasciuc, R. G. Yang

Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

CONTENTS

1. INTRODUCTION

Heat transfer at nanoscale is of importance for many nanotechnology applications [1, 2]. There are typically two types of problems. One is the management of heat generated in nanoscale devices to maintain the functionality and reliability of these devices. The other is to utilize nanostructures to manipulate the heat flow and energy conversion. Examples of the thermal management of nanodevices are the heating issues in integrated circuits [3] and in semiconductor lasers [4]. Examples in the manipulation of heat flow and energy conversion include nanostructures for thermoelectric energy conversion [5, 6], thermophotovoltaic power generation [7], and data storage [8].

Heat transfer at nanoscale may differ significantly from that in macro- and microscales. With device or structure characteristic length scales becoming comparable to the mean free path and wavelength of heat carriers (electrons, photons, phonons, and molecules), classical laws are no longer valid and new approaches must be taken to predict heat transfer at nanoscale [9, 10]. Well-known examples are the failure of Fourier law to predict the thermal conductivity of composite nanostructures such as superlattices [11, 12] and the failure of the Stefan–Boltzmann law in predicting radiation heat transfer across small gaps [13, 14]. Although much has been done in this area recently, there is still an immediate need for a better understanding of thermal phenomena in nanostructures. In addition, learning to control and manipulate heat carriers in small structures may open new paths for discovery of innovative applications.

In this chapter, we review the status and progress of theoretical and experimental investigations of thermal transport phenomena in nanostructures. In Section 2, we discuss different regimes of nanoscale heat transfer and various nanoscale heat transfer phenomena. Section 3 gives a few examples that illustrate the impacts of nanoscale thermal phenomena on modern technologies. Section 4 describes several experimental techniques developed for thermal characterization of nanostructures. Section 5 summarizes modeling tools for nanoscale heat transfer, followed by a summary of this chapter.

2. FUNDAMENTALS OF HEAT TRANSPORT AT NANOSCALE

Macroscale heat transfer is often divided into three modes: conduction, convection, and radiation [15]. Heat transfer problems are solved based on the conservation laws (mass, momentum, and energy) in combination with the constitutive equations between heat flux and temperature (or temperature gradient), for example, the Fourier law of heat conduction and the Stefan–Boltzmann law for blackbody thermal radiation,

$$q = -k\nabla T \quad \text{(Fourier law)}$$
$$q = \sigma T^4 \quad \text{(Stefan–Boltzmann law)} \tag{1}$$

where k is the thermal conductivity and σ is the Stefan–Boltzmann constant (5.67×10^{-8} W/m^2 K^4). Some distinct characteristics of heat transfer at macroscale are:

1. The Fourier law is a diffusion equation while thermal radiation between two objects separated by a nonabsorbing (or weakly absorbing) medium is ballistic.
2. Thermal conductivity is a material property which may depend on the detailed microstructure of the material but is independent of the size of the material.
3. The maximum thermal radiation heat transfer between any objects is limited by the blackbody radiation.
4. In convection, the fluid in contact with the solid assumes the same velocity and temperature as the solid at the point of contact, the so-called no-slip condition.

For heat transfer in nanostructures, some of these characteristics for macroscale heat transfer disappear. For example, heat conduction can be ballistic and similar to thermal

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (429–459)

radiation; thermal conductivity is no longer a material property; thermal radiation can be larger than the blackbody radiation; and the slip of molecules at fluid–solid interface must be considered. To understand why and when these happen, one needs to examine carefully the microscopic pictures of the heat carriers including molecules, electrons, photons, phonons (quantized lattice vibrations), or their hybrid states such as plasmons and phonon–polaritons [16, 17]. Table 1 shows some elementary properties of electrons, phonons, photons, and molecules. We will start with a discussion of the characteristic lengths of some of these energy carriers. From such a discussion, different transport regimes can be distinguished [4, 9, 10].

2.1. Characteristic Lengths and Heat Transfer Regimes

From quantum mechanics, the energy carriers have both wave and particle characteristics. At macroscale, wave phenomena such as interference and tunneling usually do not appear and we often treat the energy carriers as particles. At nanoscale, however, wave effects become important and even dominant in some cases. A key question is when one should start to consider the wave characteristics. There are a few important characteristic length and time scales that determine the answer to this question, including the mean free path, the phase coherence length, the wavelength, and the thermal (de Broglie) wavelength, which we will explain.

The mean free path is the average distance that energy carriers travel between successive collisions, such as the phonon–phonon collision in a dielectric material and the electron–phonon collision in a conductor or semiconductor. The corresponding average time between successive collisions is the relaxation time. Direct calculation of the mean free path is generally difficult, particularly for electron and phonon transports in solid. The kinetic theories and experimental conductivity data are often used to estimate the mean free path. For example, the thermal conductivities of gases, metals, and semiconductors or dielectrics can be expressed as [16–19]

$$k = \frac{1}{3} C v^2 \tau = \frac{1}{3} C v \Lambda \qquad \text{(gases)} \tag{2}$$

$$k = \frac{\pi^2 n \kappa_B^2 T}{m v_F} \Lambda \qquad \text{(electrons in metals)} \tag{3}$$

$$k = \frac{1}{3}\int_0^{\omega_{\max}} C_\omega v_\omega^2 \tau_\omega \, d\omega = \frac{\Lambda}{3}\int_0^{\omega_{\max}} C_\omega v_\omega \, d\omega \qquad \text{(phonons)} \tag{4}$$

where C is the volumetric specific heat (i.e., the specific heat per unit volume), τ is the relaxation time, v is the velocity of carriers, m is the electron mass, n is the electron number density, and v_F is the electron velocity at the Fermi surface. The integration in Eq. (4) is over all the phonon frequency and correspondingly, C_ω, v_ω, and τ_ω are the volumetric specific heat, the velocity, and the relaxation time at each frequency, respectively. This distinction is necessary because phonons are highly dispersive. If Eq. (2) is used to estimate the phonon mean free path, using the measured specific heat and the speed of sound, the mean free path can be an order of magnitude lower than that based on Eq. (4) [20–22]. In Figure 1, we show the mean free path in representative media. Also shown in this figure is an example of the phonon mean free path in silicon estimated based on Eq. (2), using the reported data on specific heat and the speed of sound, which is an order of magnitude shorter than that estimated from considering the phonon dispersion.

The phase of a wave can be destroyed during collision, which is typically the case in inelastic scattering processes, such as the electron–phonon collision and phonon–phonon collision. An inelastic scattering process is the one that involves the energy exchange between carriers. If the phase destroying scattering process occurs frequently inside the medium, the wave characteristic of carriers can be ignored and the transport falls into the diffusion regime. The measure for the phase destroying scattering events is called the phase coherence length and, for electron transport, the Thouless length [23]. Not all the scattering processes, however, destroy the phase. Elastic scattering processes such as scattering of photons by particulates and the scattering of electrons by impurities do not destroy phase. Thus, the phase coherence length is usually longer than the mean free path but is not much longer, particularly at room temperatures for electrons and phonons. Therefore we can treat them as having the same order of magnitude.

Thus, one necessary condition for the inclusion of the wave effects is that the mean free path should be comparable to or larger than the structure characteristic length, such as the thickness of a film or the diameter of a wire. This condition, however, is not sufficient for actually observing the wave effects because of three additional factors, (1) the interface scattering processes, (2) the wavelength of the carriers, and (3) the spectrum of the carriers, as will be explained.

When the wave meets an interface, it will be scattered. The most familiar example is the reflection and refraction of optical waves [24]. Electron waves and phonon waves

Table 1. Basic characteristics of energy carriers.

	Electrons	Phonons	Photons	Molecules
Propagation media	in vacuum or media	in media	in vacuum or media	in vacuum or media
Statistics	Fermi–Dirac	Bose–Einstein	Bose–Einstein	Boltzman
Frequency	0–infinite	0–Debyte cutoff	0–infinite	0–infinite
Wavelength at 300 K (nm)	~10	~1	$\sim 10^5$	~0.1
Velocity at 300 K (m/s)	$\sim 10^6$	$\sim 10^3$	$\sim 10^8$	$\sim 10^2$

Source: Adapted with permission from [10], C. L. Tien and G. Chen, *Trans. ASME, J. Heat Transf.* 116, 779 (1994). © 1994, American Society of Mechanical Engineers.

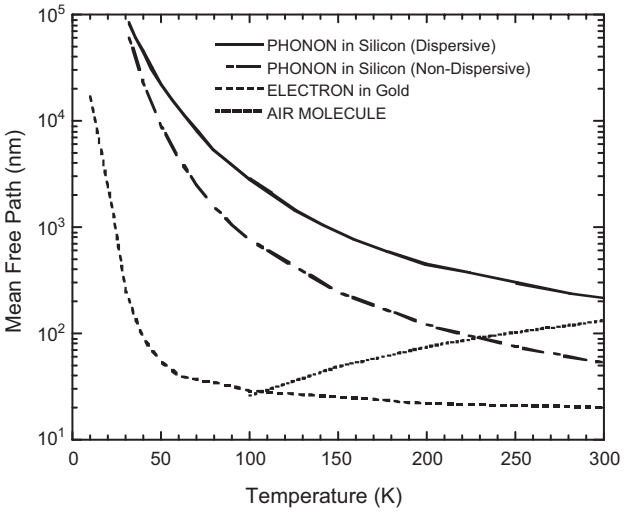


Figure 1. Mean free path for representative energy carriers.

have similar processes. For a flat interface, the phases and directions of the refracted and reflected waves are fixed relative to the incident waves. These processes thus do not destroy the phases. Periodic interface corrugations, such as surface gratings, are another example for which the incident and outgoing waves have clear fixed phase relations. Rough interfaces, however, are more complicated. If the detailed interface roughness structures are known and if the interface interaction is elastic, the directions of the reflected and refracted waves can in principle be determined from, for example, solving the Maxwell equations for photons and the Schrödinger equation for electrons and phonons. In reality, this is rarely possible and rough interface scattering is often assumed to diffuse (i.e., the reflected and transmitted waves are isotropically distributed into all directions). Often, the accompanying assumption is that the relationship between the phases of the reflected, transmitted, and incident wave is lost (i.e., the scattering is phase randomizing). Such an assumption cannot be justified easily but appears to be true in many transport processes, particularly for phonons and photons. Clearly, in addition to elastic scattering, inelastic scattering can be also strong at the interfaces and such scattering processes are phase breaking. Thus, the interface scattering can be approximated as phase breaking if the interface is rough and as phase preserving if it is smooth. Whether an interface is rough or smooth depends on the average roughness, δ, compared to the wavelength λ. For example, an approximate expression for the fraction of specular scattered phonons at an interface is [19]

$$p = \exp\left(-\frac{16\pi^3\delta^2}{\lambda^2}\right) \tag{5}$$

From this expression, we can approximately take

$$\frac{\delta}{\lambda}\begin{cases} \gg 0.1 & \text{(Rough)} \\ \ll 0.1 & \text{(Smooth)} \end{cases} \tag{6}$$

Based on the previous discussion, if the interface scattering is diffuse, the wave aspects of energy carriers can be neglected. If the interface scattering processes preserve the phase relations, coherent waves may be established over the transport domain. Even under such situations, however, wave effects may still be unobservable. There are two additional factors that one should consider. One is the wavelength of the energy carriers and the other is the thermal spread in their wavelengths.

Heat transfer usually involves a wide spectrum of energy carriers, as summarized in Table 1. However, not all the energy carriers spanning such a wide range of wavelength contribute equally to the thermal transport. The actual probability of excitation for a specific quantum mechanical state depends on the energy of the state and the temperature of the object as governed by the Fermi–Dirac distribution for electrons and the Bose–Einstein distribution for phonons and photons. The blackbody radiation is a well-known example, with the emissive power peaks at

$$\lambda T = 2898\ \mu\text{m}\cdot\text{K} \tag{7}$$

which is known as the Wien displacement law [25].

We can estimate the order of magnitude of the average wavelength of the energy carriers, λ_t, by assuming the average energy of one quantum state is $\kappa_B T/2$, where $\kappa_B (= 1.38 \times 10^{-23}$ J/K) is the Boltzmann constant and calculating the corresponding wavelength from the de Broglie relation, $\lambda = h/p$, for material waves or from the Planck relation $E = h\nu$ for phonons and photons, where $h (= 6.6 \times 10^{-34}$ J·s) is the Planck constant, p is the momentum, and ν is the frequency. This leads to

$$\lambda_t = \frac{h}{\sqrt{3m\kappa_B T}} \qquad \text{(for electrons or molecules)} \tag{8}$$

$$\lambda_t = \frac{2h\nu}{\kappa_B T} \qquad \text{(for photons and phonons)} \tag{9}$$

where v is the speed of carriers. Equations (8) and (9) are also close to the thermal de Broglie wavelength but, for accuracy, we will just call it the thermal wavelength. Figure 2 plots the wavelength as a function of temperature for these

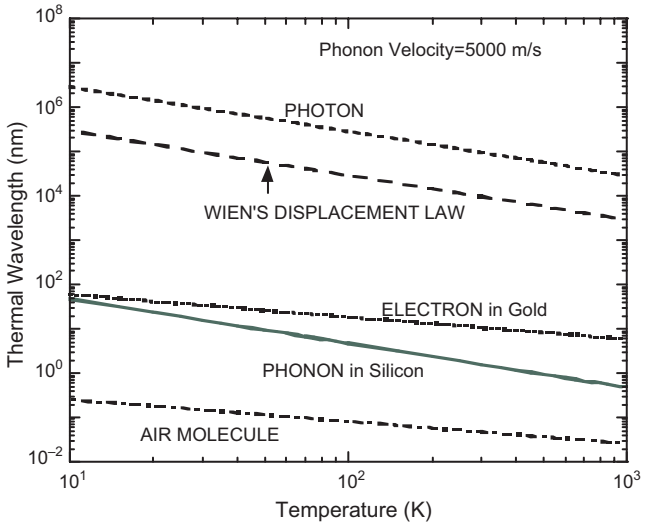


Figure 2. Thermal wavelength for representative energy carriers.

carriers with representative mass and velocity values. We also plotted the Wien displacement law. At room temperature, the wavelength orders of magnitudes are 1 Å for molecules, 10 Å for phonons, 100 Å for electrons, and 100 μm for photons.

In addition to the thermal wavelength, we should also consider the spread in energy and wavelength of the carriers. The superposition of waves forms wave packets that travel at the group velocity. In optics, the width of the wave packets is called the coherence length and is of the order of [24]

$$l_c \approx \frac{c}{\Delta\nu} \tag{10}$$

where c is the speed of light and $\Delta\nu$ is bandwidth of the radiation. This terminology will not be used here to avoid the confusion with the phase coherence length we discussed earlier. For thermal transport, the spread in thermal energy is also of the order of $\kappa_B T$. Thus, the wave packet width of the thermally excited energy carriers is of the same order of magnitude as the thermal wavelength, although a different definition can give some different results, particularly for thermal radiation [21, 26]. In the following, we will use the thermal wavelength as a measure of the wave packet size.

Because of the spread in the wavelengths of the energy carriers, the wave effects maybe smeared out and unobservable. The best example is the Young interference experiments of light passing through two slits. When the slit separation is small, interference is observable. In the opposite limit, the effect is unobservable. In this case, interference for each spectrum in the wave packet still occurs but is smeared out due to the superposition of the interference patterns of different wavelengths. In Figure 3, we show the optical transmittance of a film as a function of the nondimensional thickness (defined as the thickness parameter) obtained based on three different approaches: (1) integration of the electromagnetic wave solution over the incident spectrum, (2) optical coherence theory, and (3) ray tracing (geometrical optics) that neglects the wave effects [27]. One can see that the wave solution eventually approaches the ray-tracing approach as the film becomes much thicker than the wavelength. Thus, in the thick film limit, although consideration of the phase of the waves can lead to the correct end result, such a consideration may be unnecessary.

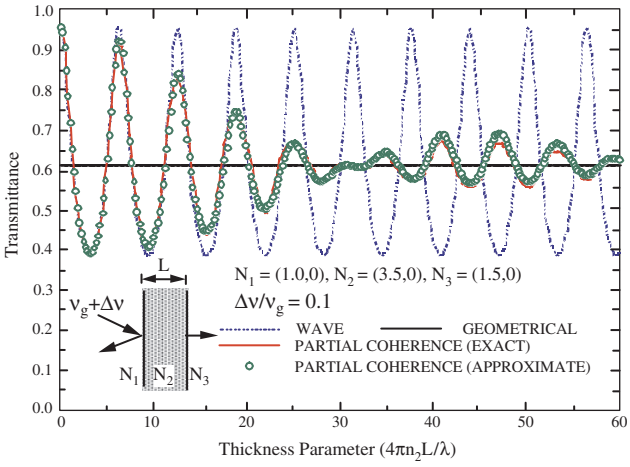


Figure 3. Transmittance of a film as a function of nondimensional thickness for a polychromatic incident wave with spectral width $\Delta\nu$ and central frequency at ν_g, calculated from three different methods, the ray tracing method (or geometrical optics) that neglects the phases of the waves, the wave method, and the partial coherence theory formulation, both including the phases of the waves. These results show that as the film becomes thicker, the wave results approach the ray tracing results. Reprinted with permission from [27], G. Chen and C. L. Tien, *Trans. ASME, J. Heat Transf.* 114, 636 (1992). © 1992, American Society of Mechanical Engineers.

Based on this discussion, we see that if the structure size is much smaller than the thermal wavelength, and if the interface scattering processes do not break the phases, the wave effects are distinct and one is likely to observe such effects as interference and diffraction. However, if the interface scattering preserves the phase but the structure is much larger than the thermal wavelength, wave effects can be obscured and not distinct. Under such conditions, as an approximation, some transport problems can be treated without considering the phase information of the energy carriers. We should emphasize, however, that this approximation is not always true. As an example of the failure of the previous approximation, we consider the acoustic wave transmission through a periodic thin film structure (i.e., a superlattice). The transmissivity was calculated with two methods, the acoustic wave transfer matrix method that includes the phases of the waves and the ray-tracing method that does not include the phase information [28]. The transmissivity is averaged thermally excited phonons incident from all allowable angles. Figure 4a and b shows the thermal conductance through a periodic superlattice, which is proportional to transmissivity. As the thickness of each layer increases, the thermal conductance approaches a constant as shown in Figure 4a. This occurs around 10 Å—the order of magnitude of the phonon thermal wavelength at room temperature. This constant thermal conductance, however, is different from that obtained with ray tracing, as shown in Figure 4b. As the number of layers in such a periodic structure increases, the figure shows a thermal conductance or transmissivity that asymptotically approaches zero as more layers are added. Clearly, in this case, ray tracing is not a good approximation to the wave approach.

Both Figure 3 and Figure 4 consider the wave propagation through thin film systems, but they lead to quite different conclusions (i.e., in the former case, the ray tracing approximation is valid while not for the latter case). We use wave packet propagation as shown in Figure 5 to explain the difference. A wave packet can experience multiple reflections in a film. In a single layer thick film (Fig. 5a), for example, the multiple reflection of a small wave packet does not overlap with itself. In a periodic structure, however, the same wave packets split at different interfaces have a chance of overlapping each other, as sketched in Figure 5b. Because of the overlapping of the same wave packets, there is always the wave interference effect, despite the fact that these effects are not reflected as periodic variations of the transmittance due to thermal averaging effects.

With the previous discussion, we can now divide the nanoscale heat transfer into several regimes as shown in Figure 6. The first demarcation line is the mean free path

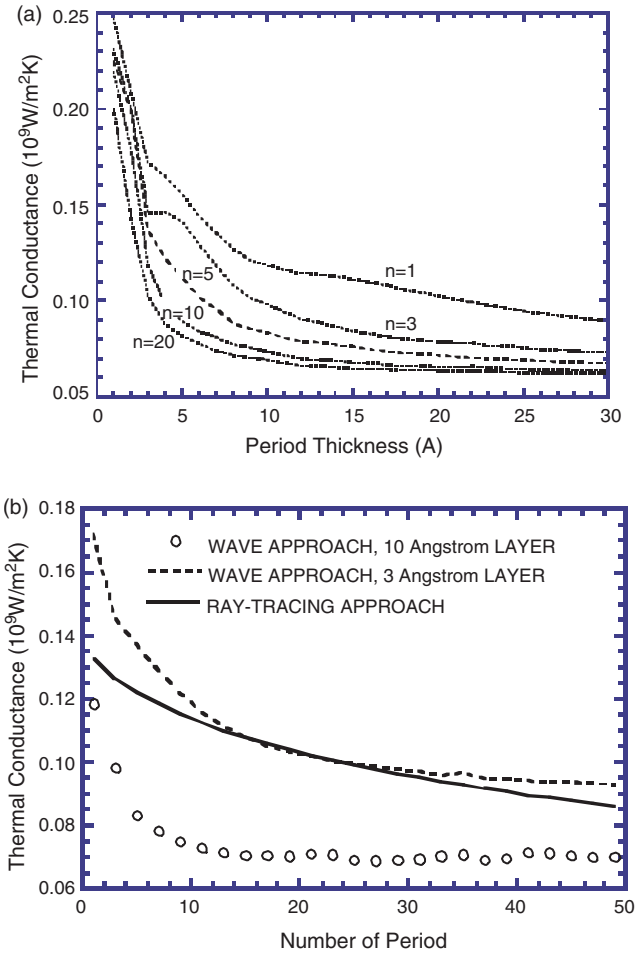


Figure 4. Thermal conductance based on calculation of acoustic waves through a periodic superlattice showing that the wave approach does not necessary lead to the same results as the ray tracing approach. Panel (a) shows that thermal conductance does not change when the layer thickness is larger than 20 Å, of the order of thermal wavelength. However, (b) shows that while the wave treatment leads to a nonzero conductance with increasing number of periods, ray tracing treatment leads to thermal conductance asymptotically approaching zero. Both calculations neglected internal scattering and assumed perfect interfaces. Reprinted with permission from [28], G. Chen, *Trans. ASME, J. Heat Transf.* 121, 945 (1999). © 1999, American Society of Mechanical Engineers.

(or phase coherence length) versus the characteristic structure size. Here, we emphasize that the mean free path and phase coherence length are values in bulk materials. Boundary effects will be discussed latter. If the structure is much larger than the mean free path, transport is a diffusion process (or drift-diffusion as for convection and electron transport under a driven current). When the structure is comparable to or smaller than the mean free path, size effects begin to appear and eventually become dominant. The actual numbers that demarcate different regimes are not to be considered accurate. For example, in rarefied gas dynamics, the continuum regime is assumed when $d/\Lambda > 100$ [29], while we used 10 in Figure 6. We further distinguish the size effects into the classical size effect regime and the quantum size effect regime, depending on whether the phase

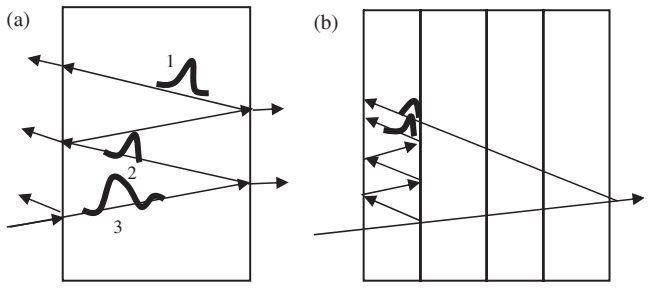


Figure 5. Traveling and interference of the same incident wave packet inside a film and a superlattice. (a) Inside a thick film, the same packet experiences multiple reflections and splitting at each reflection. When the film is thick, there is little chance that these split wave packets can overlap. The end results are that no interference beats can be observed and thus ray tracing can be used without considering the phase of the waves. (b) In a period structure, however, the same wave packets are split many times at different interfaces and it is possible a wave packet experiencing multiple reflections in one layer can overlap with that reflected into the same location from other layers. Because of this process, the ray tracing method will never lead to the same results as the wave optics even when each layer is very thick.

of the energy carriers needs to be considered. In addition to the classical and quantum size regimes, we sometimes also call these regimes particle and wave regimes or incoherent and coherent transport regimes, because it is not always necessary to treat the waves as quantized waves, such as the case for electromagnetic waves. In the classical size effects regime, one can ignore the phase of the energy carriers and trace their trajectories. Whether the transport is in the classical or quantum size effects regime depends on whether the interface scattering processes can maintain the phases of the waves. We used the terminology of interface phase-breaking scattering strength to distinguish the classical and quantum size effects regime rather than just the interface roughness, to emphasize that only complete phase breaking can justify neglecting phases associated with the carriers. In the quantum size regime, the phase information is important and must be considered. When the structure characteristic length is much smaller than the mean free path, scattering inside the medium can be neglected and the transport is ballistic. Theoretical modeling of transport in the ballistic transport

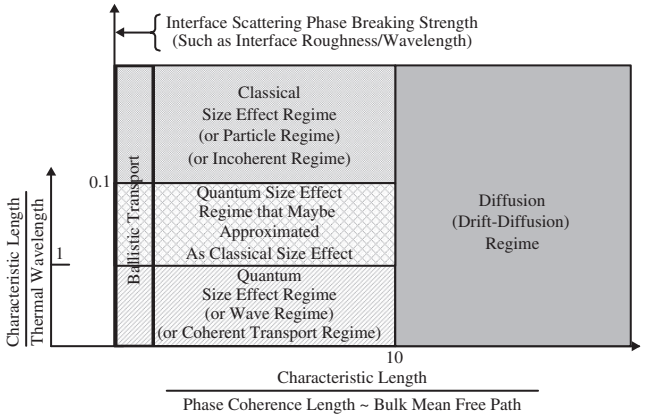


Figure 6. Different nanoscale heat transfer regimes.

regime is relatively easy. In between the ballistic and diffusive regime, both scattering at the boundary and scattering inside the medium must be considered. It is generally difficult to treat transport in this intermediate regime.

The characteristic length needs more explanation. First, the characteristic length does not have to be the structure length. As an example, in the fast heat conduction processes, a steep temperature gradient may be established over a short distance. The distance over which the temperature distribution changes significantly can be comparable to or shorter than the mean free path and in this case, size effects need to be considered. Second, a structure can have several characteristic lengths. Consider phonon heat conduction in a thin film as an example. There are two characteristic lengths. One is the film thickness and the other is the film length (or width). If the film length is much longer than the mean free path, the transport along the film plane can be described by a diffusion equation such as the Fourier law. If, at the same time, the film thickness is smaller than the mean free path, however, the thermal conductivity of the film will no longer be the same as that of the bulk constituent material of the film because lateral interfaces scatter phonons. If the lateral interface is diffuse, the transport falls into the classical size effect regime. If the lateral interface is specular, the transport falls into the quantum size regime. Thus, one can have quantized incoherent transport, quantization due to lateral interface while diffusion transport along the film plane. Electron transport in such a quantized incoherent transport regime has been studied for thermoelectric applications and it is generally agreed that the lateral quantization effect can enhance the electron energy conversion capability [30, 31]. For phonon heat conduction, however, the quantum treatment usually leads to a result that is similar to the treatment based on classical size effect treatment for a single layer thin film with specular surfaces [32] (i.e, both treatments lead to a thermal conductivity equal to that of the bulk, because the shorter thermal wavelength smears out the phonon quantization effects).

2.2. Nanoscale Heat Transfer Phenomena

With the discussion on various characteristic lengths and heat transfer regimes, we will now move on to discuss several size effects associated with three different modes of heat transfer. Because all energy carriers have both wave and particle characteristics, we will discuss first some general phenomena associated with transport in the coherent and incoherent regimes. Whether a particular heat transfer mode in a particular structure can display some of the phenomena depends on whether conditions we discussed in the previous section are satisfied.

All waves experience reflection and refraction at an interface, with the electromagnetic wave as the best-known example. For optical waves, a particularly interesting phenomenon is the total internal reflection when the refractive index of the incident medium is larger than that of the second medium and when the angle of incident is larger than the critical angle; that is [24],

$$\theta_1 \geq \theta_c = sin^{-1}\left(\frac{n_2}{n_1}\right) \tag{11}$$

where n_1 and n_2 are the refractive indices of the two media. In this case, there is an evanescent field in the second medium that does not carry a net energy into the second medium. The evanescent field decays exponentially, $e^{-x/\Delta}$, in the second medium, with a characteristic decay length

$$\Delta = \frac{\lambda_o}{2\pi n_2\sqrt{(n_1 \sin \theta_1/n_2)^2 - 1}} \tag{12}$$

Although the example is given for an electromagnetic wave, the total reflection and the evanescent wave phenomena also exist for electrons and phonons. In addition to the evanescent wave, there can also exist surface waves that decay exponentially at both sides of the interface and propagate along the interface, such as surface plasmons which are mixtures of electron and photon waves and surface phonon–polaritons which are mixed states of photons and phonons, and surface acoustic waves [33–35]. Some of the impacts of surface waves on heat transfer processes have only recently begun to be explored [36, 37].

When there are two interfaces or the interfaces have closed topology themselves, such as the two interfaces of a thin film or the surface of a sphere/wire, the waves can experience multiple reflections inside the structure. If the phase of each reflection is preserved, the superposition of the waves creates interference/diffraction effects. The new wave functions inside the structure can also be interpreted as new energy states from quantum mechanical point of view, which can also be obtained from solving the Schrödinger equation.

If the interfaces are periodic (i.e., the structures are periodic), there can be regions (for specific wavelength and periodicity combination) in which the waves are completely cancelled and no wave exists, either in specific directions or in all the directions. An example is for electromagnetic waves traveling inside a periodic thin film structure, as shown in Figure 7. In the stop band region, the reflectivity

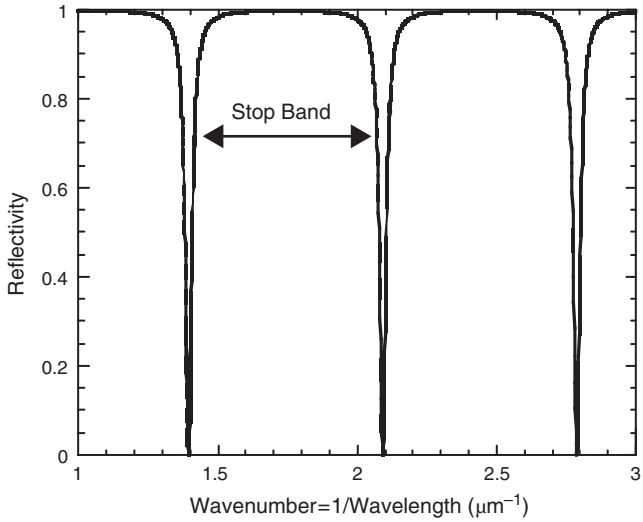


Figure 7. Calculated reflectivity of a Bragg reflector as a function of the incident photon wavelength. The refractive indices of each layer are 3 and 3.5 and the corresponding thicknesses are 417 and 352 Å for each layer, respectively.

is 100% and no waves at this frequency exist inside the film. The corresponding quantum phenomenon is the formation of electronic bandgaps that distinguish metals from semiconductors and insulators, due to the reflection of electrons in the periodic atomic potentials. Although the examples of stop bands for optical waves and that of electronic bandgap in crystalline solids have long been known, the extension of electron waves in periodic structures, such as superlattices, to form new electronic band structures [38, 39] and the extension of electromagnetic waves traveling in a three-dimensional (3D) periodic structure, to mimic what electron and phonon waves naturally do in a crystalline solids [40, 41], are relatively recent endeavors and are the precursors of the current strong drive in nanotechnology.

Wave effects offer rich possibilities because the end results depend sensitively on the phase of the waves. By changing the structures slightly, one can significantly change the phase of the waves and thus potentially engineer the transport processes. However, as discussed in the previous section, the phases are not always preserved. Phase breaking scattering events inside the solid and at the interfaces can destroy the phases and thus the wave effects. In the incoherent transport regime, the interfaces can still impact the transport because of interface scattering. These interface scattering processes generally reduce the energy flux, which can be either a problem, as for the thermal management of electronic and photonic devices [3, 4], or a blessing for other applications that demand good thermal insulation, as for thermoelectric energy conversion [6, 11].

With these general pictures, we now move on to discuss some specific heat transfer phenomena observed in nanostructures.

2.2.1. Heat Conduction

In solids, heat is conducted by electrons and phonons. In pure metals electrons dominate the conduction, while in semiconductors and insulators the dominant contribution comes from phonons. We will examine several heat conduction phenomena that are important at nanoscale, including heat conduction at a single interface, heat conduction inside thin films and nanowires, heat conduction outside nanostructures, and nonequilibrium heat conduction processes between different heat carriers.

Thermal Boundary Resistance Both electrons and phonons can be scattered at an interface. We will focus on phonons first, for which much work has been done [42]. For heat conduction perpendicular to an interface, phonon reflection implies that the energy carried by heat carriers will be reduced compared to the case when there is no interface, or equivalently, a resistance for heat flow exists at the interface. This phenomenon, called Kapitza resistance or thermal boundary resistance, has been known since the pioneering work of Kapitza [43] for liquid helium–solid interface and Little [44] for solid–solid interface, and extensive experimental and theoretical studies have been carried out in the past. At extreme low temperatures when the phonon thermal wavelength is long (Fig. 2), the interfaces are close to specular and models based on acoustic reflection and refraction for thermal boundary resistance lead to reasonably good agreement with experiment. At room temperature, however, the phonon wavelength is short and diffuse scattering can be dominant. There is no proven easy way to model the phonon reflectivity and transmissivity for such diffuse scattering processes. Two models have provided limit values for the reflectivity and transmissivity. One is the radiation model that assumes that phonons from the side with a lower maximum frequency have a unit transmissivity across the interface. By reciprocity or the principle of detailed balance, not all phonons from the side with a higher maximum phonon frequency can go over the interface. Otherwise, one side of the interface will be cooled and the other side heated, even at equilibrium. This radiation limit will provide the maximum transmissivity and thus the minimum thermal boundary resistance value. The other model, called the diffuse mismatch model, assumes that phonons are well mixed during interface scattering and there is no way to distinguish the reflected phonons from the transmitted ones on the same side (i.e., $r_{12} = t_{21}$, where r_{12} represents the phonon reflectivity for incident phonons from medium 1 into medium 2 and t_{21} is the phonon transmissivity from medium 2 into medium 1). Under this assumption, the phonon transmissivity can be expressed as [45]

$$\tau_{12} = \frac{v_2 U_2}{v_1 U_1 + v_2 U_2} \qquad \text{and} \qquad \tau_{21} = 1 - \tau_{12} \tag{13}$$

where U_1 and U_2 are the phonon energy density on media 1 and 2, respectively. This expression assumes a constant group velocity and is an extension of the low-temperature expression in [46]. If the phonon transmissivity is known, the thermal boundary resistance can be calculated from [46]

$$R = \frac{T_1 - T_2}{q} = \frac{2[1 - \langle \int_0^1 \tau_{12}(\mu_1)\, d\mu_1 + \int_0^1 \tau_{21}(\mu_2) d\mu_2 \rangle / 2]}{\int_0^1 [\int \tau_{12}(\mu_1) v_1 C_1(\omega)\, d\omega] \mu_1\, d\mu_1} \tag{14}$$

where $\mu (= \cos\theta)$ is the directional cosine and θ is the incident angle. This expression differs from the commonly used expression for thermal boundary resistance [44],

$$R_e = \frac{T_{e1} - T_{e2}}{q} = \frac{2}{\int_0^1 [\int \tau_{12}(\mu_1) v_1 C_1(\omega)\, d\omega] \mu_1\, d\mu_1} \tag{15}$$

because of the definition of temperature at the two sides of the interface [47, 48]. At the interface, phonons are in highly nonequilibrium states. The phonons coming toward the interface are at one set of temperatures (T_{e1} and T_{e2}), called the emitted phonon temperature (Fig. 8). These phonons, however, do not represent the local energy density because the reflected and transmitted phonons have different spectral distributions. If these phonons are assumed to approach equilibrium adiabatically, the final temperature T_1 and T_2 represents the local energy density. This latter temperature, defined as equivalent equilibrium temperature, is consistent with those local equilibrium temperature used in the Fourier law. At very low temperatures, as used in many experiments on thermal boundary resistance, one can place the sensors so that only T_{e1} and T_{e2} are measured [42]. At high temperatures, many experimental methods are based on extracting T_1 and T_2. For these cases, Eq. (14) should be used.

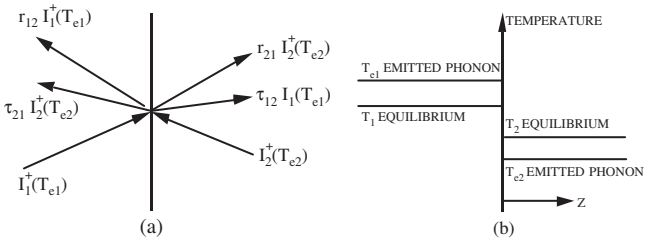


Figure 8. Illustration of phonon transport at an interface, (a) incident, reflected, and transmitted phonons, and (b) the temperature on each side of the interface. Reprinted with permission from [48], G. Chen, *Phys. Rev. B* 57, 14958 (1998). © 1998, American Physical Society.

From Eq. (14), the order of magnitude of the thermal boundary resistance is

$$R \approx \frac{2-(\tau_{12}+\tau_{21})}{C_1 v_1 \tau_{12}} \sim 10^{-9}\frac{1}{\tau_{12}} m^2K/W \qquad (16)$$

where we have used the fact that for many materials at moderate to high temperatures, $C \sim 10^6$ J/m^3 and $v \sim 10^3$ m/s. The velocity seems to be low compared to the speed of sound in many solids but is actually reasonable after considering the dispersion of the phonon spectrum. Thus, for perfect interfaces, the thermal boundary resistance is of the order of 10^{-8}–10^{-9} m^{2}K/W. This order of magnitude is consistent with molecular dynamics simulations [49] and experimental results on superlattices [50, 51].

Although this discussion is for phonons, electrons can also be reflected at an interface due to the potential mismatch at the interface. Interfacial electron transport has been studied extensively in the literature for current flow and these results can be used to predict the interface thermal boundary resistance, for example, based on the extension of the Wiedmann–Franz law from bulk to an interface [52]. If the interface is made of a metal and a dielectric, an additional energy conversion process between electrons and phonons on the metal side must be included, because eventually heat is transferred across the interface by phonons [53].

Heat Conduction in Thin Films, Superlattices, Nanowires, and Nanotubes A thin film has two interfaces. If the mean free path is longer than the film thickness, energy carriers (electrons and phonons) will be scattered more frequently at the interfaces and may experience multiple scattering. If the scattering at the interface is diffuse, as we assume for scattering due to random roughness, the thermal conductivity along the film plane will be reduced because diffuse scattering means some phonons that originally travel along the film plane direction are redirected backward. Solutions for the electrical conductivity in this classical size effect regime for a single layer of thin film, called the Fuchs–Sondheimer model, were obtained a long time ago [54–56] and are applicable to phonon heat conduction as well [3]. If the interface scattering is specular, the classical size effect model of Fuchs and Sondheimer leads to a thermal conductivity of a thin film identical to that of bulk material, because the thin film acts simply as a waveguide for the heat flow. When the film thickness is thinner than the thermal wavelength, quantum effects may be important. However, lattice dynamics calculations show that quantum effect on the energy density and group velocity leads only to a small reduction in thermal conductivity for specular interfaces [57], although there are also suggestions that the scattering mechanisms will be significantly changed, which leads to a lower thermal conductivity [58, 59]. Quite extensive studies experimentally have been carried out on the thermal conductivity of thin films. Single crystalline silicon thin films provide a model experimental system for studying the size effects, due to both the long phonon mean free path and its importance in industrial applications. In Figure 9, we show the thermal conductivity of single crystalline silicon thin films measured by several groups [60–62]. Size effects are particularly strong at low temperatures because of the increasing phonon mean free path with decreasing temperature. Polycrystalline thin films show an even larger thermal conductivity reduction, because grain boundaries greatly scatter phonons [63–70]. In the cross-plane direction, there are some measurements on the thermal conductivity of metallic and semiconducting thin films [71–73] and amorphous thin films [74–78]. Some reduction of the thermal conductivity of metallic and semiconducting thin films has been observed but the mechanisms are not well explained. For amorphous thin films, the measured thermal conductivity reduction is often explained by the inclusion of the thermal boundary resistance between the film and the substrate [74].

Superlattices are periodic thin film structures as shown in Figure 10a. Thermal properties of such structures have been studied over the last 15 years due to their importance for photonic and thermoelectric devices. Experimentally, it has been observed that superlattices have much lower thermal conductivities compared to the bulk values calculated based on the Fourier law, using the properties of their parent materials [50, 79–85]. An example of the thermal conductivity of superlattices is given in Figure 10b [83]. There are several potential explanations [11, 51]. One is the phonon spectrum change [86–91]. Figure 11a–f shows an example of the phonon spectrum of a Si/Ge-like superlattice

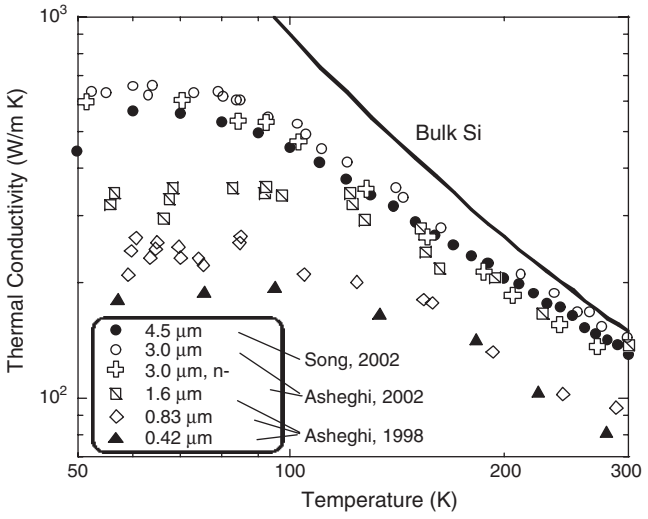


Figure 9. Thermal conductivity of single crystalline silicon thin films, showing a significant reduction as the film becomes thinner and the temperature becomes lower.

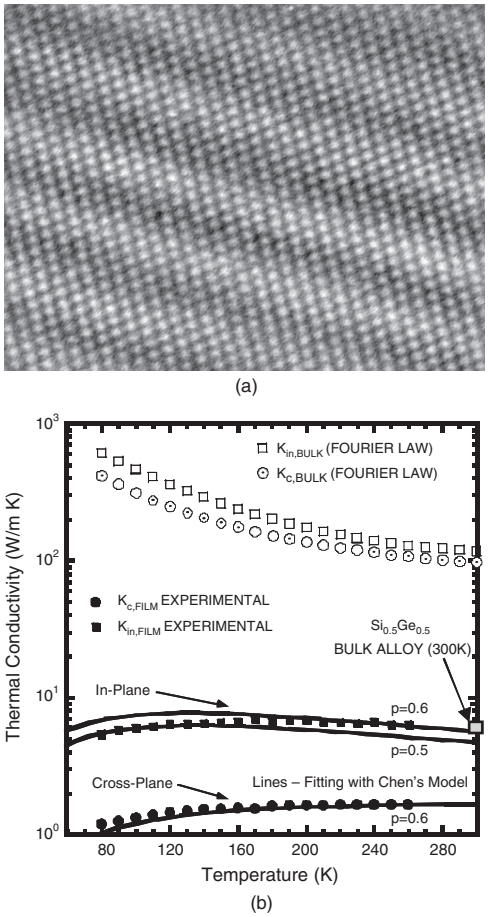


Figure 10. (a) A transmission electron micrograph of a 20 Å Si–Ge superlattices and (b) the measured thermal conductivity in both in-plane and cross-plane directions, showing a significant reduction compared to bulk values. Reprinted with permission from [83], W. L. Liu et al., *J. Nanosci. Nanotechnol.* 1, 39 (2001). © 2001, American Scientific Publishers.

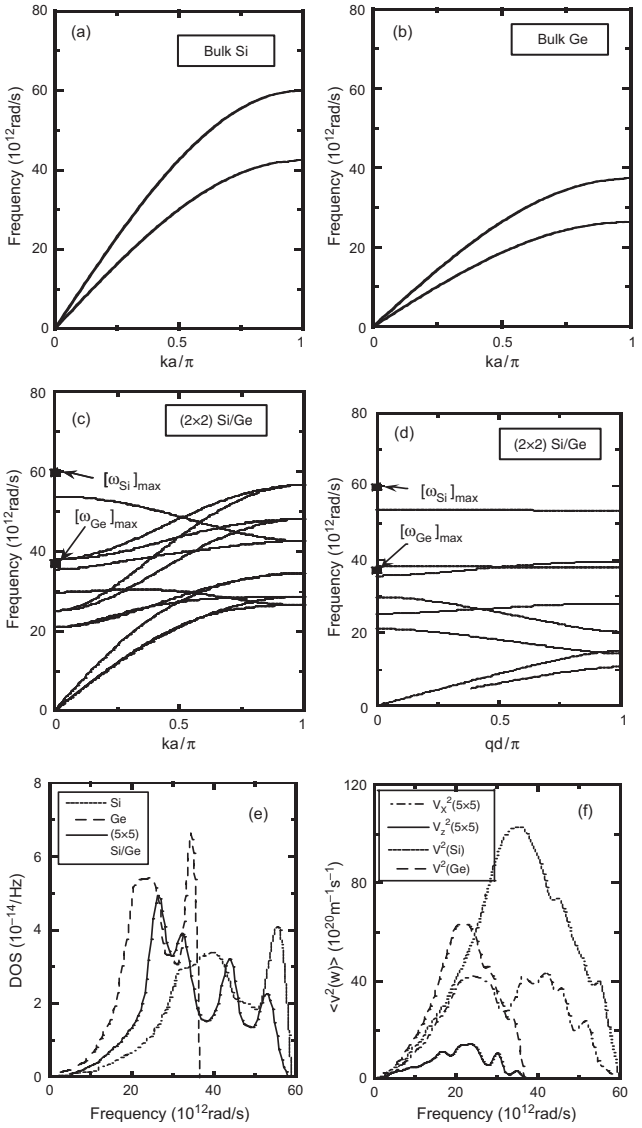


Figure 11. (a)–(d) phonon dispersion in Si, Ge, and in the in-plane (x) and cross-plane (z) direction of a Si–Ge superlattices. (e) and (f) Density of states and the group velocity. Reprinted with permission from [32], B. Yang and G. Chen, *Microscale Therm. Eng.* 5, 107 (2001). © 2001, Taylor & Francis.

calculated from lattice dynamics [32]. Compared to the phonon spectra of the bulk materials, there are several changes: (1) phonons are folded due to the new periodicity in the growth direction, (2) the acoustic phonons in the silicon layer with a frequency higher than that in the Ge layer become flat or confined due to the mismatch in the spectrum, and (3) mini bandgaps form. The impacts of these changes on the phonon density of states and group velocity are shown in Figure 10e and f. Although the change in the density of states is relatively small, there is a relatively large change in the group velocity in the cross-plane direction. This group velocity has been proposed as an explanation for the thermal conductivity reduction [89, 92]. Comparison with experimental data, however, shows that the group velocity reduction alone cannot explain the magnitude of the thermal conductivity reduction perpendicular to the film plane nor can it explain the thermal conductivity reduction along the film plane at all [32]. The reason is that the lattice dynamics model assumes phase coherence of the phonons over the whole superlattice structure and does not include the possibility of diffuse interface scattering. Models considering classical size effects lead to reasonably good agreement with experimental data [21, 48, 93]. These theoretical and experimental studies show that it is difficult to take advantage of the wave effects of phonons in phonon transport processes, primarily due to the short phonon thermal wavelength.

Thermal conductivity of nanowires is also attracting great interests [94–106]. Naturally, the interest is in for transport along the wire axis direction. Referring to the transport regime picture in Figure 6, there are several possibilities. If the surface of the wire is specular and there is no scattering inside the wire, the transport is ballistic and phonon energy states are quantized. The Landauer formalism leads to the universal phonon thermal conductance as [101, 102]

$$K = \frac{\pi^2 \kappa_B^2 T}{3h} \tag{17}$$

This universal thermal conductance has been demonstrated [101], but at temperatures <0.8 K. At slightly higher temperatures (>3 K), the diffuse interface scattering effects begin to appear [100]. If the classical size effects picture holds (i.e., if the surface is random enough to break the phonon coherence, which is likely the case at room and high temperatures, from the large amount of data on thin films and superlattices), nanowires should have even lower thermal conductivity compared to thin films due to increased surface scattering. Experimental data on single nanowires at near room temperature are just becoming available on some nanowires and indicate a significantly lower thermal conductivity, suggesting the importance of diffuse interface scattering [104]. Clearly, transport along very long wire is diffusive. A key issue is whether the lateral quantization occurs. Boltzmann equation based modeling has seen two different approaches, one based on the bulk dispersion and imposing the surfaces as boundary conditions [105, 106]. The other first assumes that boundaries create quantized phonon states, followed by the treatment of diffuse interface scattering [98]. Which of these approaches is more appropriate depends on the strength of the interfaces breaking the phase coherence. There is no consistent answer to this question at this stage. Molecular dynamics simulations may provide a way to directly compute the thermal conductivity of nanowires [105]

Carbon nanotubes are another interesting material system that differs from the nanowires. A freestanding single wall carbon nanotube has all the atoms on the surface and the phonon modes can only propagate along the axial direction [107]. This latter attribute also means that unlike a solid nanowire in which the phonon modes inside the wire can hit the boundary (assuming boundary scattering is strongly phase breaking), phonons on the carbon nanotube sheet have no boundaries to interact with. In addition, the strong modification of the phonon dispersion can also change the scattering mechanisms. It is thus suggested that carbon nanotubes can have a thermal conductivity even higher than that of diamond [108]. Measurements and simulation data so far vary widely. Molecular dynamics simulations carried out by various group also have widely varying thermal conductivity values [109–113]. Similar is the case of experimental studies. Measurements on an isolated multiwall carbon nanotubes lead to high thermal conductivity values [99] but measurements on tangled and aligned carbon nanotubes have values orders of magnitude smaller [114–116]. At this stage, the reasons for such large discrepancies are unclear. It is quite possible that although a freestanding nanotube has high thermal conductivities, a nanotube embedded in a host sees increasing interface scattering which reduces its thermal conductivity values. Continued study on this issue is important for the potential applications of carbon nanotubes.

Rarefied Phonon Heat Conduction External to Nanostructures Although the thermal conduction reduction in thin films and nanowires is a well-known phenomenon, the size effects outside for heat conduction external to nanostructures have not received much attention. Heat generated inside nanoscale regions or nanostructures eventually will be conducted to the surrounding. When the heat generation region or the size of the nanostructure is smaller than the mean free path of heat carriers in the surrounding medium, the temperature rise of the nanostructure can be much higher than that predicted by the Fourier law [117]. To see why this occurs, we consider the heat conduction surrounding a heat generating spherical region embedded inside a semi-infinite medium. The Fourier law leads to the relation between heat transfer rate, Q, and the temperature rise at the surface of the sphere, T_s,

$$Q(\text{Fourier}) = 4k\pi r(T_s - T_\infty) = \frac{4\pi}{3} rCv\Lambda(T_s - T_\infty) \quad (18)$$

where T_∞ is the temperature of the medium away from the sphere and r is the radius of the sphere, and we have used the kinetic expression, Eq. (2), for the thermal conductivity. When the diameter of the nanosphere is much smaller than the mean free path of the heat carrier in the surrounding, however, we can neglect the scattering and treat the heat transfer between the region and the surrounding as a radiation process. This approach leads to the solution for the heat transfer,

$$Q(\text{radiation}) = \pi r^2 Cv(T_{s,e} - T_\infty) \approx 2\pi r^2 Cv(T_s - T_\infty) \quad (19)$$

where we have used $T_{s,e}$ to represent the temperatures of the phonons coming out of the nanoparticle, or the emitted phonon temperatures according to Figure 8. If there is no reflection at the interface, the surface temperature as used in the Fourier law is related to the emitted phonon temperature through, $T_s = 0.5 \cdot (T_{s,e} + T_\infty)$, which leads to the second equality of Eq. (19). Comparing Eq. (19) with Eq. (18), we can see that

$$\frac{Q(\text{Fourier})}{Q(\text{Radiation})} = \frac{2\Lambda}{3r} \quad (20)$$

Thus, in the limit where the mean free path is much larger than the sphere radius, the Fourier law overpredicts the heat transfer rate. This is because the Fourier law is only applicable when there is not a large temperature gradient within one mean free path. For the previous example, the application of the Fourier law to a region much smaller than the mean free path (surrounding the sphere where temperature varies significantly) inherently implies that there is strong scattering in this region, which is only true when the mean free path is much smaller than the region, while in reality, the mean free path corresponding to the bulk thermal conductivity is much larger.

Several theoretical studies have investigated the rarefied heat conduction cases, including transient heat conduction for planar geometries [118, 119], steady state in a spherical geometry [117], and in more complicated device structures [120, 121]. Experimentally, the rarefied heat conduction effect is proposed to explain the fabrication of nanostructures by scanning tunneling microscopy [122]. A recent experiment by Sverdrup et al. [123] provides evidence of the effect. More experimental studies are needed to demonstrate this phonon rarefaction effect beyond doubt.

Nonequilibrium between Carriers In dealing with nanoscale heat transfer, it is important to identify where and how heat is generated and how heat is exchanged between different groups of heat carriers. For example, in the fast laser heating of metals, electrons usually first absorb the photon energy and relax the energy to phonons [124–126]. Depending on the electron–phonon heat exchange rate and the rate of heat input, the electrons can be heated to a much higher temperature than the phonons. Similarly, electrons under a high electric field can be heated to a temperature much higher than that of the lattice [23]. Such "hot electron" effects have been extensively studied for electronic devices. Because electrons and phonons travel at a different rate, with the electrons much faster, hot electrons can significant change the heat source distribution. For example, most heat in a MOSFET (metal-oxide-semiconductor field-effect transistor, the most important device for very-large-scale integrated circuits such as microprocessors and semiconductor memories) is generated in the drain side where hot electrons dump their energy [127]. The release of heat from electrons to phonons can also significantly affect the lattice temperature rise. For example, electrons interact more readily with optical phonons, particularly polar optical phonons as in GaAs [128, 129]. The optical phonons, however, do not carry heat as efficiently as acoustic phonons. Consequently, depending on the energy exchange rate between optical phonons and acoustic phonons, hot optical phonons can be generated. Due to the large dispersion of acoustic phonons, it is also possible that acoustic phonons are significantly out of equilibrium with each other. There are also possibilities to create cold electrons relative to the lattice temperature, which can be utilized for thermoelectric refrigeration [130]. Studies on the impacts of nonequilibrium between energy carriers and within the same type of carriers on heat transfer processes are relatively scarce.

2.2.2. Thermal Radiation

Photons can have a long mean free path as evidenced by the solar radiation traveling to the earth and can also have a very short mean free path as the rapid decay of electromagnetic waves inside a metal. The studies of radiation heat transfer thus have traditionally spanned the ballistic to the diffusive regime [25]. Because of its short thermal wavelength, most research on thermal radiation falls into the classical size effect regime (i.e., ignoring the phase of electromagnetic waves). Compared to phonons and electrons, however, the thermal wavelength of the photons is actually very long, as clearly shown in Figure 2, and it is much easier to observe wave effects for thermal radiation than in heat conduction. Thus, the most interesting aspect of radiation heat transfer in nanostructures comes from the wave aspects of thermal radiation.

Interference effects on radiative properties of thin films have been studied [131–137]. Because of the interference effects, radiation properties, such as emissivity of thin films, can be thickness dependent. Such thickness dependence may be important for emissivity control and for the temperature control of surfaces. In nonplanar geometries, diffraction effects can also significantly change the radiative properties [138]. While these effects are relatively well known, more exciting recent advancements occur in the photonic-crystal structures that are an extension of the periodic thin film structures as Bragg reflectors into 2D or 3D structures [139–142]. Similar to the electron band structure formation in periodic atomic potential, photon energy states in such periodic structures form bands and for appropriate geometries, absolute bandgap can form such that no radiation can exist in any direction. This opens the potential of emissivity control both spectrally and directionally.

Radiation tunneling is another interesting phenomenon that has been studied for its heat transfer implications [13, 14, 143–146]. From Eq. (12), we can see that an evanescent wave extends over a distance of the order of a wavelength under the total reflection condition. This evanescent wave can become propagating if a third medium is brought close to the interface between the first and second media. When this happens, total reflection no longer happens and electromagnetic waves originally confined within the first medium can deliver energy into the third medium. This phenomenon is called tunneling. Tunneling can lead to, for example, increased heat transfer between the two parallel plates, even above the radiation heat transfer between two blackbodies [14, 144].

A very interesting recent development is the exploration of surface waves such as surface plasmons and surface phonon–polaritons [36, 37, 147–150]. Surface waves exist only at specific frequencies. The surface modes decay exponentially away from the surface and are normally not useful. However, just as in the case of tunneling, the surface mode can tunnel into a third medium if it is brought into close proximity to the first interface. An example is shown in Figure 12, which shows the increased radiation heat transfer at the surface phonon–polariton frequency of boron nitride [151]. These surface modes can also be coupled by grating structures into propagating waves in free space [148]. Because the surface waves have a relatively long mean free path and because the surface modes exist only in a narrow wavelength range, the thermal emission from such surface modes can be coherent and nearly monochromatic.

These discussions focus on the photon transport. Another effect of nanostructures is to change the optical constants. For example, the energy spectrum change of electrons and phonons due to quantum size effects will create subsequent change on the optical constants. Another example is the anomalous skin effect that occurs when the electromagnetic field varies strongly over the mean free path of electrons, similar to the nonlocal phonon heat conduction effects [19, 152]. The impacts of these changes on radiative heat transfer have not been explored, except the anomalous skin effects.

2.2.3. Convection

Convection is a process concerned with heat exchange between a solid surface and a fluid. The energy transfer in convection is due to both molecular diffusion and bulk motion of the fluid in the presence of a temperature difference. As in conductive and radiative heat transfer, convection may also experience size effects. For example, the mean free path of gas molecules becomes comparable with

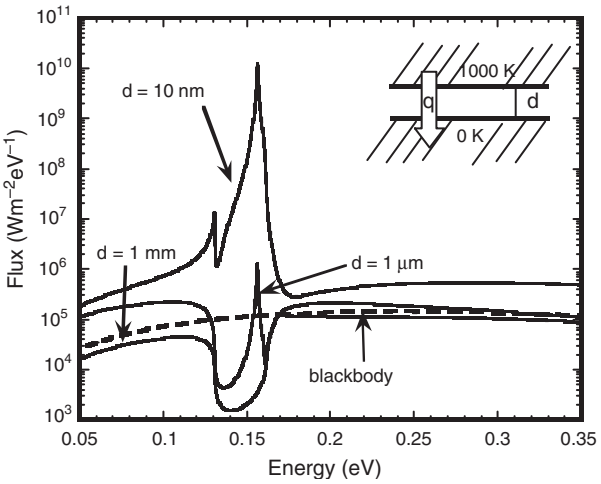


Figure 12. Spectral radiative energy transport between two half-planes of cubic boron nitride separated by a layer of vacuum of thickness d. Adapted with permission from [151], A. Narayamaswany and G. Chen, "Surface Modes for Near Field Thermophotovoltaics," unpublished.

the characteristic length of a nanostructure even at atmospheric pressures, as shown in Figure 1. This situation is similar to heat transfer in gases of low pressures and can be described by the rarefied gas theory [153]. Characteristic to this regime is the thermal creep or thermal transpiration where gas molecules creep from the cold side to the hot side along a channel with a diameter comparable to the mean free path [153]. In the case of small particles subjected to a temperature gradient this phenomenon may induce particle levitation due to thermophoresis [154]. On the other hand, liquid molecules are closely packed and have short-range interactions. Typically the continuum theory can be used to describe liquids layers as thin as 5–10 molecular diameters [155], although close to a solid surface other phenomena such as wetting, adsorption, electrokinetics, and dielectrophoresis may have to be considered [156, 157].

One way to control thermal properties of liquids is the addition of small particles to the liquid. Studies of effective thermal conductivity of solid suspensions date back to Maxwell's work on effective media theory [158]. However, recently it has been observed that the addition of nanometer size particles dramatically changes transport properties and enhances heat transfer performance of the liquids [159]. The mechanisms for enhanced heat transfer in nanofluids are not well understood [160]. Further theoretical and experimental studies are required to understand them.

3. APPLICATIONS

Energy generation and transport is a fundamental process essential to many engineering applications. Nanoscale heat transfer phenomena discussed previously have impacts on a wide range of contemporary technologies from microelectronics to bioengineering. The unusual heat transfer physics in nanostructures also opens new ways to tailor materials' properties, which in turn may lead us to innovative applications. We illustrate here with a few examples in information technologies, energy technologies, material processing, and biotechnology.

3.1. Information Technologies

Nanoscale heat transfer phenomena discussed previously have important implications for microelectronics, optoelectronics, and data storage technologies. The thermal management of microelectronics has already become a challenging issue. Nanoscale heat transfer effects may further complicate the problem [3]. Figure 1 shows that the phonon mean free path in silicon at room temperature is ~250 nm, longer than the channel length of a MOSFET device. In addition, heat is mostly generated in the drain side over a lateral dimension ~10 nm. The rarefied phonon heat conduction effect discussed before means that the Fourier law underpredicts the device temperature rise. In addition, the nonequilibrum between optical and acoustic phonons can further increase the local temperature rise. As the gate length continues to decrease, it is also predicted that the operating voltage cannot correspondingly be reduced for maintaining a low leakage current [161]. Cryogenic operation of electronic devices is being seriously considered by the microelectronics industry to improve performance. The phonon rarefaction effect will be more significant at lower operation temperatures (i.e., the deviations from the predictions of the Fourier law will become larger), because of the increasing phonon mean free path in bulk materials with decreasing temperature. How this effect will impact the scaling of microelectronic devices has not been carefully examined.

Another key component of current technological applications is semiconductor diode lasers which are usually made of multilayer thin films. Two major types of semiconductor diode lasers, edge–emitting lasers and vertical-cavity surface emitting lasers (VCSELs), are currently commercialized and used in a wide range of applications including laser printers and high density optical disks, playing a central role in today's optical communications. The excessive temperature rise in VCSELs, partly due to the high electrical resistance created by the heterointerfaces and partly due to the low thermal conductivity of the multilayer structures, had been the road block for their commercialization until a few years ago [4]. Studies on the thermal conductivity of superlattices were partly motivated by this issue. Although this problem has partly been solved for some low power applications, heating is still the most important factor limiting the power output and is more severe for near infrared lasers due to the large nonradiative recombination that is eventually converted into heat [162–165]. Edge emitting devices have less of a heating problem than VCSELs but facet damage is often associated with heating [166]. The high power outputs required to achieve reliable optical performance coupled with the low thermal conductivity of multilayer thin film structure often result in a short device life [4, 167]. The reduced thermal conductivity calls for careful design of the devices to minimize the number of interfaces and to allow efficient heat removal from the gain region.

While heating in integrated circuit and in semiconductor lasers is undesirable and needs to be "managed," it is actually used extensively as a means for writing data onto storage medium such as CDs and magnetic hard drives. For

magnetic hard drives, one of the major factors that determine data storage density is the separation between the slider and the magnetic disk known as the flying height. In present-day hard drives the flying height is around 50 nm while the disk rotates at $\sim$10, 000 rotations per minute. The airflow between the slider and the disk is crucial in maintaining the height of the disk and is very different from the prediction of continuum fluid mechanics at this small spacing [168]. Another factor to be taken into account is the thermal creep [168], and this effect has also been proposed for novel levitation-based recording systems [169]. Equally important to the storage density is the size of the magnetic domain. It is difficult to increase the storage capacity of current magnetic disks beyond a density of 40 Gb/in^2 [170] because of the superparamagnetic effect that affects magnetic domain stability. New magnetic materials with larger magnetic anisotropy are being developed [171] and heating has been proposed as a means to assist the writing processes [8, 172, 173]. Meanwhile, alternative technologies are being developed. Some rely again on heating, such as thermomechanical data storage, where data bits are written on polymer substrates by heated atomic force microscope (AFM) tips. A storage density as high as 400 Gb/in^2 has been demonstrated [8]. The writing process as well as the data reading requires highly localized spatial and temporal heat deposition. Many rewritable CDs are based on the phase change caused by laser heating [174–176]. For such applications, it is desirable to limit the heating within a small domain. Nanoscale heat transfer effects including reduced thermal conductivity of thin films and nonlocal heat conduction surrounding nanoscale heating spots can be utilized to confine heat for writing smaller spots.

These examples emphasize the small length scales involved in nanodevices and nanomaterials. A short time scale is also becoming increasingly important. Similar questions can be raised for transport at short time scales as for the small length scales. Lasers can deliver a pulse as short as a few femtoseconds (1 fs $= 10^{-15}$ s). Energy transduction mechanisms at such short time scales can differ significantly from that at macroscale [126]. Microelectronic devices are pushing to the tens of gigahertz clock frequency with a much shorter transient time. The device temperature rise in such short time scales can be very different from predictions of the Fourier law.

3.2. Energy Conversion

Energy conversion is a field that may greatly benefit from nanoscale energy transport phenomena. An example is thermoelectric cooling and power generation based on the Peltier effect and the Seebeck effect, respectively [177]. The efficiency of a thermoelectric device is determined by the thermoelectric figure-of-merit of the material $ZT = \sigma S^2 T/k$, where S is the Seebeck coefficient, σ is electrical conductivity, k is the thermal conductivity, and T is the absolute temperature. Currently, the best commercially available thermoelectric materials based on Bi_2Te_3 and its alloys has a ZT of $\sim$1. Thermoelectric devices based on this material cannot compete in performance with other well-established technologies for energy conversion. Several approaches have been explored to increase the thermoelectric figure-of-merit [5, 178]. Among those, low-dimensional thermoelectric materials structures such as quantum wells, superlattices, and nanowires have been extensively investigated [179-183]. Quantum size effects on electrons in nanostructures can be utilized to improve the electron energy conversion capability [30], while the reduced thermal conductivity of nanostructures due to increased scattering of heat carriers from interfaces can be exploited in reducing the denominator of ZT. In the past two years, different low-dimensional thermoelectric materials that show improved cooling capacity as compared to bulk materials have been reported [185, 186]. This suggests that phonon engineering is a primary tool in developing better thermoelectric materials.

The wave effects on radiation heat transfer can be utilized for thermophotovoltaic energy conversion in different ways. Photonic crystals can have the spectral emissivity tuned to match the bandgap of semiconductor photovoltaic cells [187]. Tunneling effects are being explored to increase the energy density delivered from the heat source to the photovoltaic cells [13, 188]. Surface waves, such as surface plasmons and surface phonons, can be utilized for nearly monochromatic thermal radiation, which increases the efficiency of thermophotovoltaic energy conversion [36, 151, 189] as well as the energy density if small gaps are employed to take advantage of the high energy density of surface waves.

3.3. Nanomaterial Synthesis and Nanofabrication

Nanomaterial synthesis and nanostructure fabrication are two areas where thermal transport is often important but has received less attention. Size-dependent melting point depression of nanoscrystals may be used to develop techniques for fabrication of cheap electronics [190]. A different example is related to multilayer structure masks currently used in extreme ultraviolet lithography and X-ray lithography [191]. The heat generation in these structures and its impact on the structural fabrication accuracy has yet to be investigated. Ultrafast laser nanomachining [192] and nanoimprint lithography [193], both based on short time interaction of a laser pulse with sample surface, also call for better understanding of flow and heat transfer at nanoscale and short time scale. The synthesis of nanoscale materials is a wide-open field and many nanomaterial and nanostructure synthesis methods being developed raise intriguing nanoscale heat transfer questions. For example, nanowires and carbon nanotubes have been made with different methods such as chemical or physical vapor deposition, filling of templates through liquid phase, vapor deposition, and electrodeposition [194–198]. Understanding the transport processes during nanomaterial formation will allow better control of the final material quality. Scanning probe based fabrication methods also often involve nanoscale heat transfer issues that are not well understood. For example by focusing a laser beam onto an AFM or scanning tunneling microscope tip, structures with dimensions down to 10 nm can be fabricated [199, 200]. In a different experiment an infrared laser is focused on an AFM tip

in contact with a PMMA substrate to produce pits of several hundreds Å in size [201]. Understanding the interaction between electromagnetic radiation and nanostructures and its coupling, the heat conduction, is obviously very relevant for improving these fabrication methods.

3.4. Biotechnology

Nanoscale heat transfer is also making in-roads in biotechnology. Inductively heating gold nanoparticles attached to DNA strands has been found to be a promising method to control DNA hybridization [202]. The heat transfer mechanism between the nanoparticle and DNA strand has yet to be elucidated.

4. EXPERIMENTAL TOOLS

There are two key experimental directions in studying nanoscale heat transfer phenomena. One is the thermophysical property measurements of nanostructures such as thin films and nanowires. The other is the temperature measurements, particularly temperature mapping surrounding nanoscale devices. In this section, we briefly describe some of the techniques used for thermophysical properties characterization and temperature measurement of nanometer size samples. Additional details can be found in reviews on similar topics available in the literature [11, 64, 203–205].

4.1. Thermophysical Properties Characterization

Measuring thermophysical properties is always a challenging task, even at macroscale. The thermophysical property characterization of nanostructures and microstructures is even more challenging. The measurements of the thermal conductivity, for example, normally require the determination of heat flux and temperature drop between two points of the sample, which becomes much more difficult in nanostructures than in macrostructures. Taking thin films as an example, the measurement of thermal conductivity perpendicular to the film plane (cross-plane direction) requires the determination of the temperature drop over the film thickness that ranges from nanometers to micrometers. Because thin films are often deposited on substrates, it is difficult even to create a reasonable temperature drop across the film without creating a large temperature rise in the substrate, not to mention determining the temperature drop across the film. The thermal conductivity measurement parallel to the film plane (in-plane direction) may look easier because temperature sensors can be placed along different locations in the film surface. However, the substrate makes difficult the determination of actual heat flow in the film. To overcome these difficulties, different strategies have been developed for measuring thin film thermal properties in different directions. In the cross-plane direction, the general strategies are (1) creating a large heat flux across the film while minimizing the heat flux in the substrate and (2) measuring the surface temperature rather the temperature drop. These can be realized by, for example, using microfabricated heaters directly deposited on the film. By using a small heater width, the heat flux is large when going through the film but the heat spreading inside the substrate significantly reduces the temperature drop in the substrate. Another example is to use transient or modulated heating that can limit the heat affected region close to the film. In the in-plane direction, one often used strategy is to remove the substrate such that the heat flux through the film can be uniquely determined.

In the following, we will first discuss thin film thermophysical property measurements and divide our discussion based on different heating and temperature sensing methods: the electrical heating and sensing based on microheaters and sensors, the optical heating and sensing methods, and the combined electrical/optical methods. After these discussions, we will review several methods used for nanowire property measurements.

4.1.1. Electrical Heating and Sensing Based on Microheaters and Microsensors

Microheaters and sensors can be fabricated onto the samples using microlithography techniques developed in the semiconductor industry. In some methods, the heaters are also used directly as the temperature sensor while in other techniques, separate heaters and temperature sensors are used. The advantages of the electrical heating and sensing methods, compared to optical heating and sensing, are that the amount of heat input into the sample can be precisely controlled and the temperature rise can be accurately determined. By using microheaters and microsensors, high heat flux can be created and temperature rise can be pinpointed at micrometer scales, depending on the sensor dimension. We will further divide our discussion for measurements in the cross-plane and in-plane directions.

Cross-Plane Thermal Conductivity Measurement of Thin Films One major method for measuring the cross-plane thermal conductivity of thin films is the 3ω method, initially developed for measuring the thermal conductivity of bulk materials and later extended to thin films [74, 206]. Thermal conductivity of thin films down to 20 nm in thickness has been measured using this technique [74, 206, 207]. The method has also been adapted for simultaneous measurement of the in-plane and cross-plane thermal conductivity of anisotropic thin films [83] and freestanding membranes [208]. In the 3ω method a thin metallic wire is deposited onto the sample surface to act as both a heater and a temperature sensor. An ac current with angular modulation frequency (ω) passing through the wire causes Joule heating and ac temperature oscillations of the wire with a frequency of 2ω. Due to the temperature dependence of the electrical conductivity, the resistance also has a 2ω component proportional to the ac temperature rise. The voltage drop along the wire thus contains a third harmonic (3ω) component that depends on the ac temperature rise of the heater. Since the heater measures only the temperature rise on the front surface (T_f), the temperature rise at the interface between the film and the substrate (T_s) must be determined. This can be done by two approaches. One method, often referred to as the slope method, is to calculate the temperature rise at the film–substrate interface based on the thermal conductivity of the substrate. The latter can be determined

from the frequency dependence of the measured front surface temperature [206] because the temperature drop across the film itself is frequency independent when the film is thin. Figure 13a shows an example of measured temperature rise at the front surface and inferred substrate temperature.

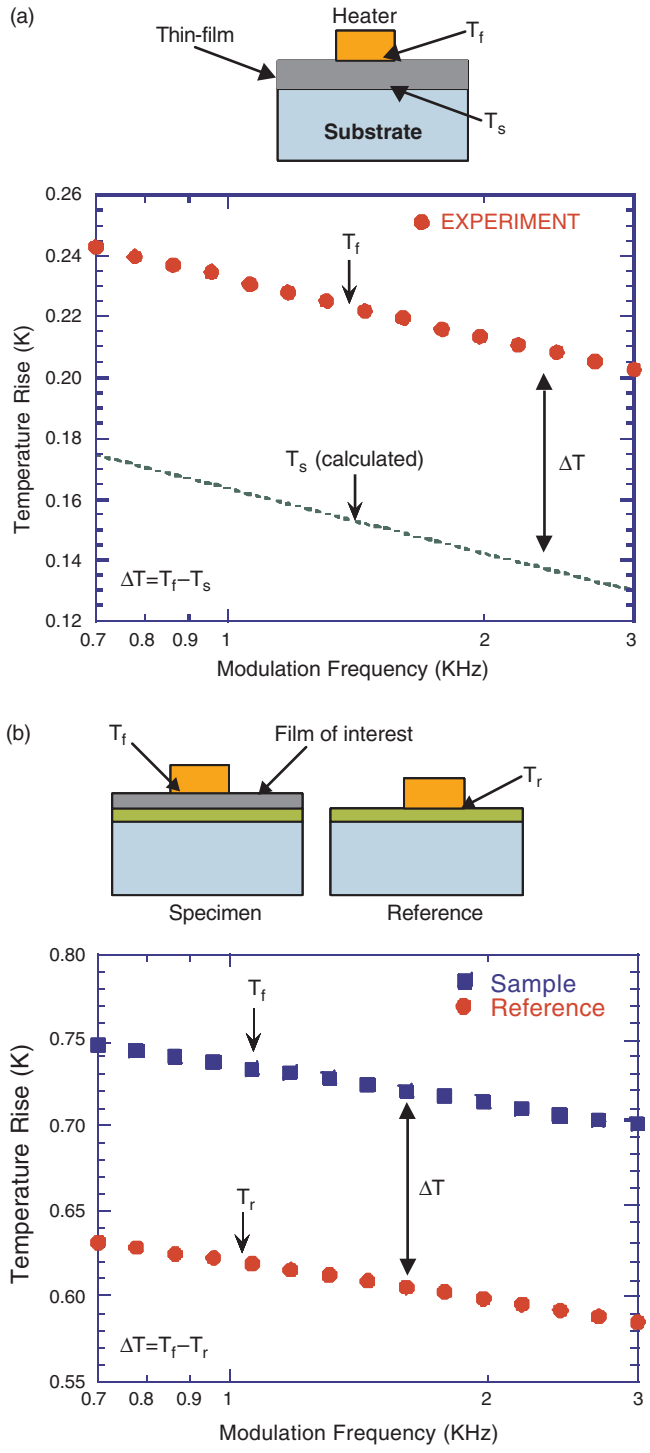


Figure 13. (a) Experimental temperature rise (T_f) and calculated substrate temperature (T_s) in the slope method. (b) Experimental temperature rise for the specimen (T_f) and reference sample (T_r) in the differential method.

Detailed modeling, however, shows that the determination of the substrate thermal conductivity based on the front side temperature measurement is prone to error when the film is relatively thick. A different way to estimate the temperature drop across the thin film is to measure it experimentally by a differential technique [209], which measures the difference of the front surface temperature rise between the actual sample and a reference sample that does not have the film. An example of the temperature rise measured by the heater onto the specimen (T_f) and reference sample (T_r) is shown in Figure 13b. Usually, the heater width is chosen to be much wider than the film thickness such that heat conduction across the film can be treated as one-dimensional (i.e., in the cross-plane direction only). Although simplified one-dimensional steady-state models are often used in extracting the cross-plane thermal conductivity of the thin film, it is important to carry out more detailed modeling when the contrast in properties between the film and the substrate becomes small. By choosing heaters of different width, for example, through the use of two heaters, one with a width much larger than the film thickness such that it is most sensitive to the cross-plane thermal conductivity, and the other with a width comparable to the thickness such that lateral spreading effect is important, both in-plane and cross-plane thermal conductivity of the thin films can be determined [83, 209].

The modulation technique as used in the 3ω method has several advantages. One is that the ac temperature field can be controlled by the modulation frequency. With reasonably high modulation frequency, the ac temperature field is confined to the region close to the heater such that the substrate can be treated as semi-infinite. This avoids the influence of the boundary condition at the substrate side. Another advantage of the modulation technique is that the ac signal is less sensitive to the radiation heat loss compared to dc measurement methods. The ac modulation also leads to the possibility of determining the substrate properties in addition to the film properties.

In addition to the 3ω method, a multisensor steady-state method was developed to measure the thermal conductivity of SiO_2 thin films deposited on silicon [210, 211]. This method relies on the temperature measurements by adjacent sensors to back up the temperature underneath the heater. Because the measurement is done at the steady state, the boundary condition at the back side of the substrate is important for determining the temperature underneath the heater.

In-Plane Thermal Conductivity Measurements Several in-plane thermal conductivity measurements have been developed by making freestanding thin film structures [65, 212, 213]. One of the configurations is to fabricate heater and temperature sensors on a large membrane as schematically shown in Figure 14a. Heat generated from the heater spread from the membrane to the frame. The other configuration shown in Figure 14b is to further shape the membrane into a narrow line and to pass current directly through the line, if the membrane is conducting or semiconducting. In this case, heat spread out along the line axis direction and the average temperature of the line is measured and correlated to the thermal conductivity.

Depending on the thermal conductivity of the film to be measured, there are several considerations one should pay

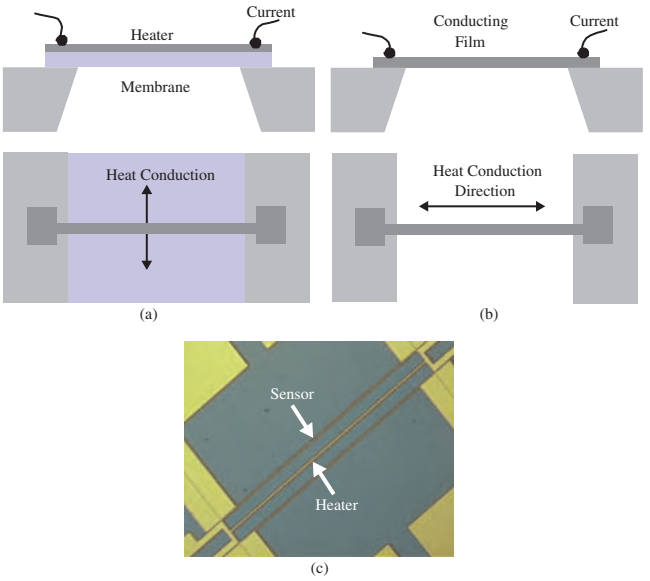


Figure 14. In-plane thermal conductivity measurement based on (a) membrane and (b) bridge structures. A fabricated membrane with the heater and the temperature sensor is shown in (c).

attention to. If the membrane method is applied to measure the thermal conductivity of a low thermal conductivity film, the structure should be designed such that heat conduction is one-dimensional, otherwise two-dimensional simulation must be carried out to back up the thermal conductivity of the film. For low thermal conductivity films, heat leakage through the heater itself may be a significant component of the total heat transfer, and thermal radiation and convection loss can further increase the uncertainties in determining the thermal conductivity. Sometimes, two different membranes are used to determine the thermal conductivity, by assuming that the effective heat transfer coefficients for both geometries are the same [212, 214]. This assumption is dangerous if the experiment is done in air because natural heat transfer coefficient scales with the heated area. Ac- or transient-based heating and measurements can reduce the effects of heat loss through the metal heater and by thermal radiation [215]. If the measured membrane has a relatively high thermal conductivity, the spreading thermal resistance in the substrate supporting the membrane can be important. Multiple temperature sensors sometimes are fabricated along the membrane to eliminate this uncertainty but the additional loss along these temperature sensors should be considered [216]. A better location is to place the temperature sensor on the substrate at the place where the membrane meets the substrate, as shown in Figure 14c.

The bridge method constrains the heat flow to one dimension [65, 212, 217]. In this sense, it avoids the potential complication of the two-dimensional heat conduction effects in the membrane. However, the thermal conductivity depends on temperature profile along the bridge, which further depends on the radiation and convection heat losses and on the thermal resistance at the two ends of the bridge. In addition, the method is only applicable if a current can pass through the film (or another conducting film with known properties is deposited onto the film).

In addition to removing the substrate, we also mention the 3ω method to measure the in-plane thermal conductivity by exploring the heat spreading effect with small heaters [83, 209, 218]. When a high thermal conductivity film, such as silicon, is separated from the substrate by a low thermal conductivity layer, such as SiO_2 [22], surface temperature can also be sensitive to the in-plane thermal conductivity [76].

4.1.2. Photothermal Methods

Microheater and microsensor based methods require microfabrication facilities. In comparison, optical based methods are more flexible and are noncontact techniques. However, since it is difficult to determine accurately the amount of heat absorbed by the sample, the optical heating based methods usually do not lead directly to thermal conductivity. Instead, these methods measure the thermal diffusivity or thermal effusivity of the sample under investigation by monitoring the sample response to time dependent heat input at the boundaries. Pulse or modulated heating is often used. By using short pulses or high frequency modulated heating, the heat affected region can be confined to close to the film such that excessive temperature rise in the substrate can be avoided.

A variety of methods have been developed based on modulated laser beam heating of the sample. The modulation heating is favored because it allows phase-locked detection of the temperature-induced signals. In these experiments, a modulated laser beam (pump) is focused onto the sample surface, producing a local rise in temperature. The detection of temperature increase can be done by various means. For example, in the photothermal reflectance method, the sample reflectivity is monitored by another laser beam (the probe) focused on same location with the pump beam [219]. The photothermal displacement technique monitors the physical displacement of another laser beam due to thermal expansion of the sample [220]. In the photoacoustic methods the surface temperature is detected by measuring the acoustic signals coupled to the piezoelectric detectors either through the air on the front surface or through the substrate to the back surface [221]. In the photothermal radiometry method, the blackbody radiation from the sample surface is detected [222]. Another way to detect temperature increase is to measure the deflection of a laser beam parallel to the sample surface, caused by the gradient of the refractive index of the gas layer adjacent to the heated spot (Mirage method) [223]. The photothermal reflectance method can also be employed to visualize the thermal wave on the surface of the sample, by moving one beam with respect to the other [224, 225]. Typically the pump beam is scanned and the probe beam remains fixed in order to avoid surface roughness effects on the reflected beam. This technique may reveal defects or grain boundaries [225].

The modulation frequency in the techniques described is typically limited to a few megahertz and often below 100 kHz. Under such modulation frequency, the ac temperature field penetrates into the substrate. Interpretation of the photothermal signals thus involves the substrate thermophycial properties. Although the latter may also be determined from the measured signal, the inclusion of the substrate thermophysical properties as unknowns also can

cause more uncertainty in the film thermophysical property determination. A method to avoid the effects of the substrate is the transient photothermal technique [226], often called pump-and-probe technique, where the transient heating pulse ranges from nanosecond to femtosecond. The main advantage of using pulse lasers is that, for a particular pulse regime, the penetration depth of the thermal wave may be small compared with film thickness. Therefore heat diffusion at nanometer length scale can be measured directly. When picosecond or femtosecond pulse lasers are used, it is not possible to detect such fast changes in temperature with typical measurement instrumentation. In this case, the temporal response is measured by delaying the probe pulse with respect to the repetitive pump pulse and measuring its reflectivity for various delay times. Other experimental setups of this technique use nanosecond lasers [227]. In this case the temperature decay is still measurable by fast radiation detectors.

Photothermal methods have been used for measuring thermophysical properties of a wide range of materials such as metallic, semiconductor, or dielectric films [219], and superlattices [80] as well as for studying grain boundaries in semiconductors [228], interface thermal contact resistance [227, 229], and thermal properties colloid metal particles [230].

4.1.3. Mixed Optical/Electrical Heating and Detection Methods

There are also several methods that employ optical heating and sensor detection, or electrical heating and optical detection. Ac calorimetry is an example where optical heating is combined with thermocouple detection, if the response time requirement is not high, as in the measurement of low thermal diffusivity samples [204]. Thermal diffusivity of the sample is determined by measuring the temperature rise at different distances from the heated spot using fine thermocouples [204] or resistive thermometers [231]. Under appropriate conditions, the thermal diffusivity along the film plane can be calculated from phase or amplitude data. Cross-plane thermal diffusivity measurement is also possible by detecting the temperature rise at the back side of the sample with a thin resistive thermometer [231]. A different way to detect the thermal response of optically heated samples is to use a fast thermoelectric effect. In photothermoelectric technique, the specimens are heated on one side by laser radiation or white light collected from a high power lamp [232, 233]. The temperature is measured on the other side by a fast responding thermocouple, which consists of the contact point between a sharp wire and the electrically conductive sample surface. The technique may be employed in thermal diffusivity measurement of freestanding films or thin films on substrates transparent to the heating beam. In addition the light spot can be scanned around the detection point to image the thermal wave on sample surface. In addition to optical heating, electrical heating is also combined with optical detection to determine the thermal diffusivity of thin films. For example, electrical heating is combined with thermal-reflectance change or thermal emission to determine the thermal diffusivity of thin film [227, 234].

4.1.4. Nanowires and Carbon Nanotubes

While the experimental methods for thermal characterization of two-dimensional structures are relatively mature, only a few attempts have been made toward thermal property characterization of one-dimensional structures. Earlier experiments on thermal transport in one-dimensional structures have been carried out by Potts and co-workers [95–97]. The motivation of the work was to observe phonon quantization in nanowires made of GaAs single crystal that is known to have a large phonon mean free path. Freestanding arrays of highly doped GaAs are resistively heated, as in the bridge method for thin film thermal conductivity measurement. Their experiment, done at cryogenic temperatures, shows that electrons and phonons are at nonequilibrium and the measured thermal conductance is due to the electron conduction. Related to the nonequilibrium interactions are the experiments of Seyler and Wybourne on the electrical conductivity of metallic nanowires on silicon substrates [94]. They observed local peaks in the electrical resistance as the dc field applied across the wires was increased. The results were explained considering the electron–phonon interaction and the quantization of the acoustic phonon in the small wires. A different testing structure, using nanowires as the suspension for a heated island, was used by Schwab and coworkers [101]. They observed quantized conductance at 1 K. At slightly higher temperatures, their experimental data can only be explained by considering some diffuse scattering at the interfaces.

A different method for thermal conductivity characterization of one-dimensional structures was developed more recently by Shi and co-workers [99, 235]. Their idea is to use freestanding microfabricated structures (islands) whose temperature can be controlled and measured individually by thin-film resistors deposited onto the surface. The nanowire is then suspended between two islands and the thermal conductivity measurements are carried out by applying Joule heating to one of the islands, while the temperatures of both islands are monitored. Key challenges of this method include the minimization of the thermal contact resistance between the specimen and the microfabricated structure, and the maximization of the thermal resistance of the suspension structure. This method has been employed to measure the thermal conductivity of multiwalled carbon nanotubes and Si nanowires with various diameters [104].

The key challenge for the property measurement of nanowires lies in sample preparation, especially aligning electrodes to nanowires or nanowires to fabricated testing structures. So far, existing reports have relied on sophisticated micro- and nanofabrication techniques. There is a need for continued development of measurement techniques and more systematic study on the thermophysical properties of nanowires.

4.2. Temperature Measurement

In addition to thermophysical property characterization, measurements of temperature distributions are also of interest in many emerging technologies. The infrared thermal imaging techniques are limited to a resolution of a few micrometers, far from the resolution needed for thermal mapping in surrounding nanostructures. Several techniques

for thermal mapping at nanoscale have been proposed and are currently under development.

4.2.1. Scanning Thermal Microscopy

Scanning probe microscopes have provided nanoscale topographical resolution and it thus appears natural to locate a temperature sensor at the apex of the probes for thermal mapping. In fact, the first attempts in thermal mapping, after the invention of scanning tunneling microscope but before the birth of the atomic force microscope, was to use a thermocouple tip to measure surface topology of nonconducting surfaces according to conductance variation [236]. Several strategies have been proposed to place temperature sensors at the apex of AFM tips [237–239]. One method is to use a thermocouple as the cantilever and the other method is to place a thermistor at the apex. Both methods have gone through many improvements in the thermal sensor fabrication [240]. The estimated spatial resolution is 30–50 nm. One key issue with contact thermal mapping is the topology effect. Surface roughness changes the thermal contact resistance between the sample and the sensing tip. Low thermal conductance cantilevers have helped to minimize the heat flow from the sample to the cantilever, which reduces the topology effects on thermal imaging [235]. Also important to the accuracy of temperature estimation are the heat transfer mechanisms between the sample and the sharp tip [235]. Thermal images from scanning thermal microscopy (SThM) have provided new information for understanding the heat generation and heat conduction mechanisms in nanostructures. However, AFM based thermal imaging still has not reached the quantitative stage. An interesting theoretical question is what the resolution of SThM means since thermal transport around the SThM can be highly ballistic and nonequilibrium.

4.2.2. Near-Field Thermal Imaging

The near-field scanning optical microscope has the capability to break the optical diffraction limit [241]. In near-field infrared microscopy [242] similar probes have been used to obtain thermal images of 2 μm wide metal strips. The aperture diameter was of 1.5 μm, approximately one order of magnitude smaller than the dominant wavelength of thermal radiation at room temperature. A method to further increase the spatial resolution of the temperature measurement is to use visible radiation to probe the local change of reflectivity with temperature [243]. However, while the spatial resolution of near-field microscopy depends primarily on the aperture size, the low transmittance through the aperture probes drastically limits measurement sensitivity. This is even worse for thermal emission measurements due to the small thermal signal. One way to overcome this problem is to use solid-immersion lenses that could improve the spatial resolution without loss in sensitivity [244]. So far, near field thermal imaging has not reached the same level as SThM in terms of the image quality and resolution.

4.2.3. Thermal Reflectance Imaging

The photothermal method relies on the change of reflectance with temperature. This effect has been used extensively in the property measurements. Because the temperature dependence of the reflectance is generally small, $(10^{-4}–10^{-6})$/K, phase-locked detection schemes are often used, which requires periodic heating [224, 246]. For photothermal reflectance based property measurements, the periodic heating is realized by an external light source [224, 245]. For the thermal imaging of electronic devices, the devices are often biased with an ac current to periodically vary the temperature fields. Although such ac biasing may not correspond to the actual device operating conditions, the thermal mapping obtained leads to qualitative information on the hot spots. Transient heating has also been used, leading to an absolute measurement of the temperature distribution [247]. The thermal reflectance method can have a resolution of the order the laser wavelength used.

4.2.4. Raman and Luminescence Based Temperature Mapping

Raman spectroscopy measures the Stokes and the anti-Stokes shift of the laser line due to the phonon scattering of light. The intensity of the Stokes and the anti-Stokes lines depends on the phonon population and thus the temperature. From the intensity ratio of the two lines, the phonon temperature can be determined. Alternatively, the linewidth can also be used for temperature determination. Raman-based temperature mapping has been used in the investigation of the facet heating of semiconductor lasers [166, 248–250]. Similarly, photoluminescence peaks also depend on temperature and have been used for temperature mapping [251]. The advantage of the Raman spectroscopy method is that the ratio of the Stokes and the anti-Stokes line intensity depends on the phonon temperature only. However, the best temperature resolution reported so far is of the order of a few degrees.

5. ANALYTICAL TOOLS

The analysis and modeling of nanoscale heat transfer phenomena are challenging and complicated by several factors. One is that the mechanisms of nanoscale heat transfer are not well understood. Figure 6 shows different possible regimes of heat transfer. The determination of which regime a specific problem at hand falls into is nontrivial. Another major problem is that heat transfer rarely occurs only at nanoscale because eventually heat is dissipated into the macroscopic environment. Many heat transfer problems are multidimensional and multiscale spanning from nanoscale to macroscale. Tools for analyzing different problems are still evolving. In this section, we will divide our discussion into four directions: (1) purely ballistic transport, (2) classical size effects, (3) quantum size effects, and (4) multiscale modeling.

5.1. Purely Ballistic Transport

If all the geometrical length scales involved in the transport processes are much smaller than the energy carrier mean free path, the internal scattering of heat carriers can be completely neglected (i.e., transport is purely ballistic). An example is the radiation heat transfer in a vacuum enclosure, which is a topic in standard heat transfer textbooks.

The short mean free paths of electrons and phonons mean that this situation only occurs in nanostructures and very often at extremely low temperatures when the mean free path is long. Modeling of heat transfer in this purely ballistic limit is relatively simple. In thermal radiation, the concept of radiation shape factors is often used for heat transfer calculations. Typical radiation heat transfer treatment is limited to the incoherent regime where the phase information can be neglected. A more general formulation is the Landauer formalism that treats transport as a transmission process [252]. The energy transfer between two points in space is

$$\begin{aligned} q_{12} &= \frac{1}{8\pi^3}\sum_p \Big[\int E(k_1)v(k_1)\tau_{12}(k_1)f_0(T_{e1}, E_{fe1})d^3k_1 \\ &\quad - \int E(k_2)v(k_2)\tau_{21}(k_2)f_0(T_{e2}, E_{fe2})d^3k_2\Big] \\ &= \frac{1}{8\pi^3}\sum_p \Big[\int E(k_1)v(k_1)\tau_{12}(k_1)[f_0(T_{e1}, E_{fe1}) \\ &\quad - f_0(T_{e2}, E_{fe2})]d^3k_1\Big] \end{aligned} \tag{21}$$

where E is the energy per carrier, v is the carrier velocity, τ_{12} is the transmissivity from point 1 to point 2, k_1 are the wavevectors at point 1, f_0 is the distribution function, and the summation is over all polarizations of the heat carriers. We have used T_{e1}, E_{fe1} to emphasize that f_o represents the carriers leaving the point 1, not necessarily the local equilibrium temperature, as we discussed in Figure 8 for the thermal boundary resistance. The second equality of the Landauer formalism for heat flux comes from the principle of the detailed balance, which says that at equilibrium, there is no net flux.

Under the Landauer formalism, the key is to calculate the carrier transmissivity from point 1 to point 2. Both the wave and particle descriptions can be included under the Landauer formalism. In Section 2.1.1, for example, we discussed several models used to evaluate the phonon transmissivity across an interface and from the Landauer formalism, one can get the expression for the thermal boundary resistance of an interface expressed by Eq. (15). Similarly, the classical thermal radiation calculation based on the radiation shape factors can be understood as the calculation of the transmissivity. Landauer formalism can also be applied to obtain the universal thermal conductance of a nanowire, by considering the quatization of the energy spectrum of nanostructures [101]. Through the calculation of the transmissivity of acoustic waves and electromagnetic waves in layered structures, the Landauer formulation has been used to study the thermal conductance of superlattices and for radiation heat transfer across small spaces and for phonon heat conduction in thin films [28].

5.2. Classical Size Effects Regime

In Figure 6, we defined the classical size regime as the region when the mean free path is long compared to the characteristic length and when the scattering at the boundaries is strongly phase-breaking and randomizing. Under this situation, the energy spectrum of the heat carriers is close to that in bulk materials, with the complication that boundaries alter the trajectory of the heat carriers. We also grouped part of the quantum size regime as approximately treatable with the approaches for classical size effect described in this section. This occurs when the thermal wavelength is shorter than the domain size, but as explained in Section 2.1, this approximation does not always work and should be used prudently. In the classical size regime, the major new physics in addition to transport in bulk materials is the boundary scattering. This is a conceptually relatively easy addition. Commonly used tools for dealing with such size effects are the Boltzmann equation and Monte Carlo simulation.

5.2.1. Boltzmann Equation Approach

The Boltzmann transport equation (BTE) applies to dilute particles such as gas molecules, electrons, phonons, and photons. In its general form, the BTE can be written as [19]

$$\frac{\partial f}{\partial t} + \mathbf{v}\cdot\nabla_{\mathbf{r}} f + \mathbf{F}\cdot\nabla_{\mathbf{p}} f = \left(\frac{\partial f}{\partial t}\right)_c \tag{22}$$

where f is the statistical distribution function of an ensemble of carriers, which depends on time t, position vector $\mathbf{r}$, and momentum vector $\mathbf{p}$. $\mathbf{F}$ is the force vector applied to the particles. The key to the Boltzmann equation is the scattering term, which is the term that restores the system to equilibrium. Quantum mechanical principles are often used to deal with scattering. The perturbation treatment in quantum mechanics leads to the Fermi golden rule of calculating the scattering probability from one quantum state to another. A general expression of the scattering integral can be formally written based on the scattering probability and the distribution function [19]. This leads to an integral-differential form of the Boltzmann equation which is difficult to solve but has often been treated in thermal radiation transport in the form of equation of radiative transfer [25, 253]. For phonon and electron transport, as well as gas transport, the relaxation time approximation is often used,

$$\left(\frac{\partial f}{\partial t}\right)_c = \frac{f_0 - f}{\tau(\mathbf{r}, \mathbf{p})} \tag{23}$$

where f_0 is the equilibrium distribution (i.e., Fermi–Dirac distribution for electrons, Bose–Einstein distribution for phonons, and Maxwell–Boltzmann distribution for gas molecules) and $\tau(\mathbf{r}, \mathbf{p})$ is the relaxation time as a function of position and momentum and is often approximated as energy-dependent rather than wavevector-dependent [254]. When multiple scattering coexists, for example, carrier–carrier scattering and carrier–impurity scattering, the Mathessien rule is often used to obtain the total relaxation time [19].

There are two different approaches to deal with the boundary scattering of nanostructures on transport processes. One is to add an extra boundary scattering term into the relaxation time through the Mathiessen rule [19]. This approach is simple and has been widely used in dealing with low temperature transport problems in bulk materials where size effects are important, but it is not accurate because while carrier–carrier scattering and carrier–impurity scattering are volumetric processes, interface scattering occurs at

the surfaces [255]. The other approach is to treat interfaces and boundaries through boundary conditions to the Boltzmann equation. In this latter approach, the relaxation time in bulk materials without any size effects is first determined. The effects of the interfaces and boundaries of nanostructures are imposed as the boundary conditions.

Classical size effects on transport processes and properties of all the energy carriers have been studied in the past. The classical size effects on electron transport were summarized in the book of Tellier and Tosser [56]. Flow and heat transfer of rarefied gas has also been studied extensively in the past [256]. Size effects on phonon transport started in the 1930s through Casimir's pioneering work [257]. The last two decades have seen a surge of investigations on phonon size effects in low-dimensional structures [11, 12].

The most studied phonon size effect is the thermal conductivity reduction in thin films. Phonon heat conduction in thin silicon thin films has been measured and modeled by Volklein and co-workers on polycrystalline thin films for thermoelectric applications [258, 259] and by Goodson and Ju for single crystalline silicon thin films [3, 22]. The modeling for the in-plane thermal conductivity of single crystalline thin films is a simple extension of the Fuchs and Sonderheim classical treatment of the electron size effects [67, 260, 261]. For polycrystalline films, the gray boundary scattering adds another relaxation time that has been grouped into the bulk relaxation time according to the Mathiessen rule. A model based on geometrical consideration rather than solving the Boltzmann equation was developed by Flik and Tien [262]. Heat conduction perpendicular to a thin film was dealt with in several papers [46, 260–265]. Majumdar [263] solved the phonon Boltzmann equation assuming that the two surfaces of the film are black phonon emitters. Chen and Tien [260] developed a model for the thermal conductivity of quantum wells based on the BTE for both the in-plane and the cross-plane directions. The in-plane thermal conductivity is based on Fuchs' solution [54] while the cross-plane direction is based on the approximate solution developed for the photon transfer equation.

Size effects in multilayer structures are more difficult to model than a single layer. The incoherent particle transport models for thermal transport in superlattices based on the Boltzmann transport equation were developed by Chen and co-workers in a series of papers for transport in both in-plane and cross-plane directions [21, 48, 93] and by Hyldgaard and Mahan for the in-plane direction [20]. The models by Chen and co-workers assume partially diffuse and partially specular interfaces, with the fraction of specular interface scattering left as a fitting parameter. Figure 15a and b shows examples of the model fitting with experimental data for both in-plane and cross-plane directions. The experimental data are from Yao [79], Yu et al. [266], and Lee et al. [50]. The good agreement seems to indicate that the phase-breaking scattering processes dominate.

The Boltzmann equation approach as has been used in thin films and superlattices can be readily extended to one-dimensional structures, such as nanowires [102, 105]. Walkauskas et al. [106] solved the Boltzmann equation for square nanowires and compared the thermal conductivity of nanowires to thin films. As expected, additional lateral surfaces cause a more pronounced reduction in the thermal

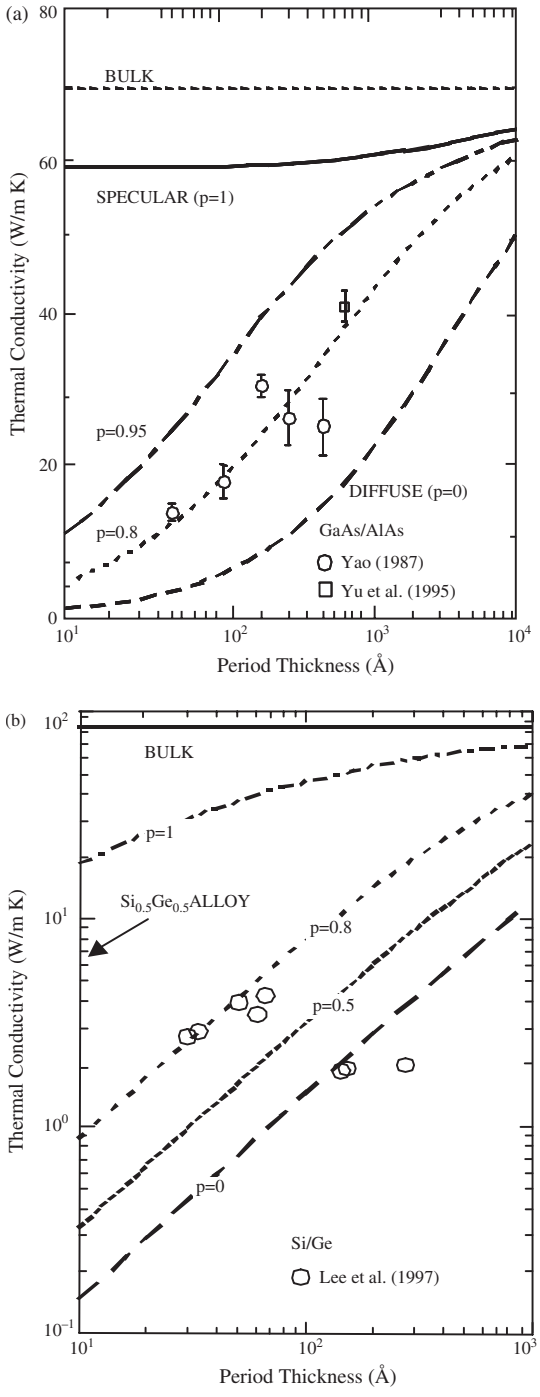


Figure 15. Modeled thermal conductivity of superlattices. (a) In-plane for GaAs–AlAs supperlattices. Reprinted with permission from [21], G. Chen, *Trans. ASME, J. Heat Transf.* 119, 220 (1997). © 1997, American Society of Mechanical Engineers. (b) Cross-plane for Si–Ge superlattices based on the Boltzmann equation. Reprinted with permission from [48], G. Chen, *Phys. Rev. B* **57**, 14958 (1998). © 1998, American Physical Society. The experimental data by Yao [79], Yu et al. [266], and Lee et al. [50] are also plotted.

conductivity of wires compared to films. Nanowire superlattices have also been modeled recently [45].

There have been many modeling studies in the past on the thermal conductivity of porous materials due to their applications in thermal insulation [267]. These models, however, cannot be applied to nanoporous materials because they did not include size effects, as is clearly required from experimental data [268]. Several attempts to include the size effects have been made in the modeling of the thermal conductivity of nanoporous silicon [269–271] and opals [272]. Size effects on nanocomposite thermal conductivity have been considered based on the inclusion of thermal boundary resistance [273]. One major difficulty in the Boltzmann equation approach lies in the solution of the equation. The Boltzmann equation is a phase space equation with seven variables. Direct numerical solution is generally difficult. There have been a few treatments applying the Boltzmann equation to nonplanar geometries that are more relevant to actual device configurations [117, 274]. More detailed discussion will be given in Section 5.4.

5.2.2. Monte Carlo Simulation

Monte Carlo simulation solves the Boltzmann transport equation in a statistical framework. Over the last decade or so, there has been tremendous advancement in the development of Monte Carlo solution techniques for the radiative transfer equation and the Boltzmann equation for electrons and holes in semiconductors [275–280]. Only a few reports using Monte Carlo technique for phonon transport have been published in the past. Peterson [281] employed the Monte Carlo method to simulate phonon transport in a confined space while Klitsner et al. [282] performed Monte Carlo simulations to obtain temperature distribution in a crystal. More recently Mazumder and Majumdar [283] considered phonon dispersion as well as various phonon scattering mechanisms to study heat transport in complex geometries and to predict the thermal conductivities. The result showed that by fitting one parameter using experimental data of thermal conductivity at one temperature, predictions of the thermal conductivity of silicon agree well experimental data over wide temperature range. In addition, it was able to capture both the ballistic and diffusive limits. The major advantage of this method is that the simulation code can be easily adapted to complex geometries. In addition, Monte Carlo simulation can separately treat various scattering mechanisms. The major disadvantage of the Monte Carlo method is its slow speed.

5.3. Quantum Size Effect Regime with Scattering

While it is relatively straightforward to treat wave effects if no scattering exists, the physical picture becomes less clear and mathematical treatment becomes more complicated when both wave effects and scattering coexist. If there is no scattering and the detailed geometries and atomic compositions of the nanostructures, particularly the interface structures, are known, the energy states of the nanostructures can be solved from the Schrödinger equation. Examples are the electron and phonon energy states of a perfect thin film or quantum well. The existence of internal scattering, however, usually distorts the energy states from these ideal calculations. Imperfect interface structures, which are usually unknown, further add to the complexity. The quantum form of the Boltzmann equation does exist [284]; similarly, quantum Green's function methods may also be applicable for treating nanoscale heat conduction problems [285]. None of these approaches, however, have been implemented so far for realistic nanostructures. So far, all the treatments for situations when both wave and scattering are in existence are all based on various assumptions that may be valid under certain conditions. There are two approximation strategies. One approach, for example, is to assume that the quantized energy states form first, usually calculated under perfect interface conditions. These quantized states are then treated as particles based on the Boltzmann equation. We will call this quantized incoherent transport. Clearly, whether this approximation is valid or not depends on the strength of the boundary scattering in breaking or randomizing the phases.

In this section, we will discuss several approaches that include, to a certain extent, the wave mechanisms, including (1) quantized incoherent transport, (2) the molecular dynamics simulation, and (3) the fluctuation–dissipation theorem approach.

5.3.1. Quantized Incoherent Transport

The quantized incoherent transport first computes the energy spectrum of the nanostructures. The subsequent motion of these quantized energy carriers, however, is treated as incoherent. Representation treatments include electron transport in quantum wells and quantum wires for thermoelectric energy transport studies [30, 178]. Electron energy states are quantized due to lateral confinement. Transport along the film plane or wire axis, however, is diffusive, either due to interface scattering or internal scattering. Experimental measurements of the electrical conductivity of nanowires seem to validate such an approach [286, 287].

For phonon heat conduction in thin films, nanowires, and superlattices, a similar approach has also been taken [58, 59]. The phonon spectra of nanowires and thin films have been computed based on continuum acoustic wave equations or the lattice dynamics method. For heat conduction along the film plane or wire axis, the assumption of a perfect interface does not lead to much change in the thermal conductivity compared to bulk materials [57]. This is also anticipated from a pure particle-based approach because in these cases, the structures simply act as waveguides [54, 55]. Only when diffuse interface scattering is included can these approaches lead to thermal conductivity predictions that are comparable to experimental observations. However, it is not clear at this stage whether such quantized states do form or not, and if yes, as some Raman spectroscopy would suggest [288] there is no quantitative information on the fraction of the quantized states compared to the bulk states.

Phonon heat conduction in superlattices offers another test to the assumption of a quantized incoherent transport model. Different lattice dynamics models have been used to compute the phonon spectra in various superlattices and these spectra are used to calculate the phonon group

velocity [88–90]. The Boltzmann equation is used to calculate the thermal conductivity using the new group velocity and the phonon density of states, assuming that the relaxation time of the superlattices is equal to that of its parent materials. These attempts, however, have not lead to quantitative agreement with experimental results, and in fact, even the qualitative agreement is poor as the lattice dynamics would lead to the thermal conductivity decreasing with period and eventually to a flat thermal conductivity that does not change with period. Figure 16 compares the lattice dynamics result with experimental results and clearly shows that neither the absolute values nor the trends are consistent with experimental data in both in-plane and cross-plane directions [51], except in the very thin period limit where the slight recovery of experimental thermal conductivity in some superlattices seems to agree with the trends of lattice dynamics predictions. In comparison, the incoherent transport models based on the assumption that phase coherence is not important leads to results that are in agreement with experimental data, particularly in the thick period regime [48, 93], as shown in Figure 15. This does seem to suggest that the phonon heat conduction processes in thin films are dominated by incoherent transport.

Why do bulk crystal solids form phonon band structures that can be used in the Boltzmann equation calculation, while the phonon band structure-based calculation for ideal superlattices seems to fail to explain the thermal conductivity? To answer this question, we can use the wave method to calculate how many layers are required to form the superlattice spectra that assume an infinite number of layers. Transfer matrix based calculation shows that 10 periods will suffice. Thus, phonons must remain phase coherent over such periods to form new superlattice bands. If diffuse interface scattering randomizes the phase, the calculated phonon spectrum based on ideal interfaces is not valid and no superlattice mode or only partial modes form. In addition, internal scattering can also destroy the phases because 10 periods can be much longer than the phonon mean free path.

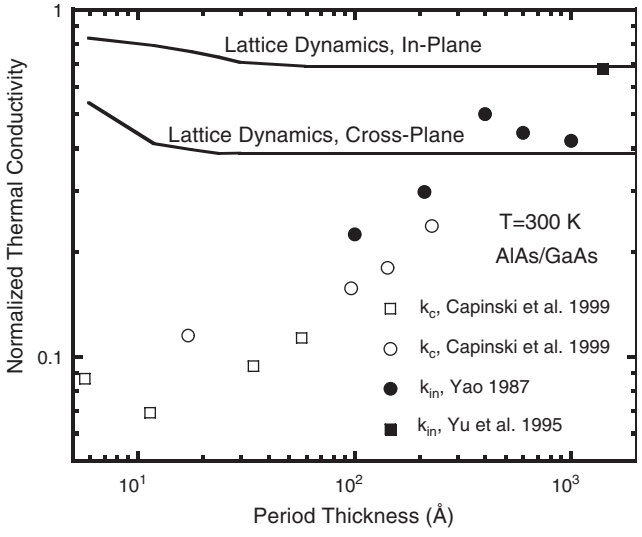


Figure 16. Comparison of lattice dynamics with experimental data. The experimental data are from Yao [79], Yu et al. [266], and Capinski et al. [80]. Adapted with permission from [51], B. Yang and G. Chen, in "Chemistry, Physics, and Materials Science for Thermoelectric Materials: Beyond Bismuth Telluride" (M. G. Kanatzidis, T. P. Hogan, and S. D. Mahanti, Eds.). Kluwer Academic/Plenum, New York, 2003. © 2003, Kluwer Academic/Plenum.

One way to treat both the quantized and the bulk energy states is to introduce an imaginary wavevector into the Bloch wave functions. When the imaginary wavevector is large, the waves are highly damped and only sample few unit cells of the original crystals making up the nanostructrure and thus tend to recover to the bulk dispersion. When the imaginary wavevector is zero, the lattice waves extend over the whole structure and the new energy bands will form. The method was first used by Pendry for electron transport [289]. Simkin and Mahan introduced it to lattice dynamics for calculating superlattice thermal conductivity [290]. Their model is able to predict a minimum in the cross-plane thermal conductivity and the period dependence. However, the Simkin and Mahan model does not consider the diffusive interface scattering, which is probably one major cause of the reduced mean free path in superlattices. As a consequence, their model cannot explain the in-plane thermal conductivity reduction in superlattices, nor the absolute magnitude of thermal conductivity reduction in the cross-plane direction. Yang and Chen related the imaginary wavevector to the interface diffuse scattering [51, 291]. This approach enables the calculation of the thermal conductivity of superlattices in both in-plane and cross-plane directions, both leading to good agreement with experimental data, as shown in Figure 17.

5.3.2. Molecular Dynamics Simulation

Like lattice dynamics, classical molecular dynamics (MD) simulation also solves Newton's second law to track the motion of all atoms in a system [292–294]. The difference between lattice dynamics and molecular dynamics is that lattice dynamics uses harmonic forces while molecular dynamics inputs the true potential of atomic interaction. Bolstered by advances in computation power, MD has become a powerful tool for studying material formation processes and material properties and it is increasingly employed to study heat transfer [295–306]. Its key limitations are determination of accurate potentials and the size of systems that can be studied. Fortunately well-tested empirical potential are available for most of the material of interest. For example the Stillinger–Webber potential is often used for silicon [307]. System size is currently the greatest limitation. Recent large-scale MD simulations of fluid flows involve ~200,000 atoms and solids of 10–100 million atoms depending on necessary times of simulation [308, 309]. Unfortunately, the mean free path of heat-carrying phonons in silicon crystals is ~3000 Å [3, 48]; yet even a 1000 Å on a side cube of silicon contains ~50 million atoms and requires a long run time to accumulate phonon statistics.

Equilibrium and nonequilibrium MD simulations have been used in heat transfer studies. In the nonequilibrium approach, a temperature difference or energy flux is imposed by altering the atomic dynamics in localized boundary regions [310, 311]. In general, nonequilibrium techniques suffer from three major drawbacks: (1) the simulation

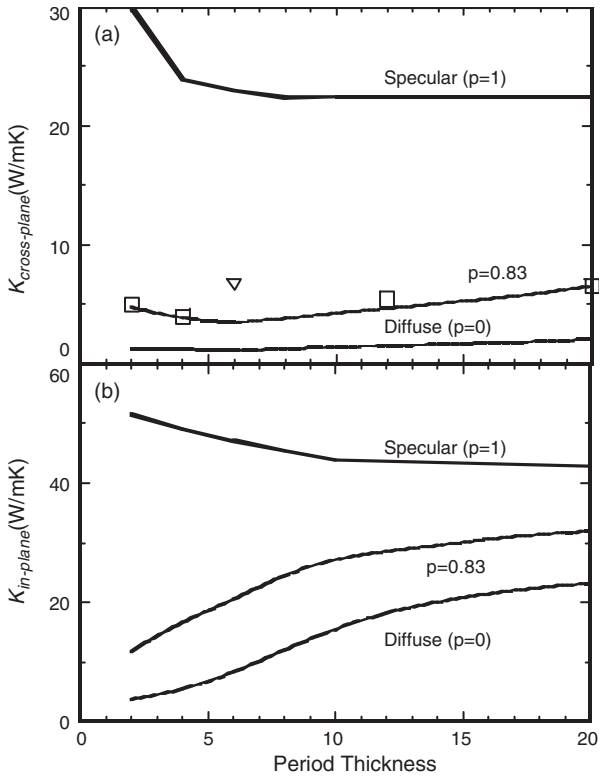


Figure 17. Thermal conductivity of GaAs/AlAs superlattices obtained from incoporating diffuse interface scattering into a lattice dynamics model; experimental data from Capinski et al. [80] are also plotted. Adapted with permission from [51], B. Yang and G. Chen, in "Chemistry, Physics, and Materials Science for Thermoelectric Materials: Beyond Bismuth Telluride" (M. G. Kanatzidis, T. P. Hogan, and S. D. Mahanti, Eds.). Kluwer Academic/Plenum, New York, 2003. © 2003, Kluwer Academic/Plenum.

domain must contain enough atoms in the boundary regions to stabilize the boundary temperatures, which increases the overall cost of the simulation; (2) the systems that can be simulated are smaller than the phonon mean free path and their small size also restricts the maximum phonon wavelength; and (3) an unphysically large temperature gradient is required to converge temperature statistics, making it difficult to determine the thermal conductivity at a particular temperature. So far, the nonequilibrium MD simulation techniques have been applied to thin films, nanowires, and carbon nanotubes [112, 301, 312].

The equilibrium approach relies on small statistical temperature fluctuations to drive instantaneous heat fluxes [313–315]. Although the time average of this heat flux vanishes, the thermal conductivity of the system to the disturbance can be calculated from the instantaneous heat flux autocorrelation function according to the Green–Kubo formula [314]. Convergence is slow, but the equilibrium approach is free of the three drawbacks associated with the nonequilibrium technique (periodic boundary conditions effectively remove the mean free path limitation [301, 303]), although care still must be taken to avoid artificial size effects of the finite simulation domain. Equilibrium methods have been used to diagnose the thermal conductivity of many materials including recent computations of high thermal conductivity materials such as silicon and diamond [301, 302]. The MD simulation of thermal conductivity of superlattices has also been explored [316–319]. Recently, Daly and co-workers used MD simulation to investigate the effects of interface mixing on the thermal conductivity in GaAs/AlAs superlattices [70]. Their simulations show that for perfect interface, the results are similar to that of lattice dynamics. For the case of superlattices with interface mixing, the simulation results are much closer to experimental data, further reinforcing the discussion we had previously on the heat conduction mechanisms in superlattices.

From the time history of individual atoms in a MD simulation, phonon spectral characteristics can be analyzed by discrete Fourier methods. Thus the relaxation time and other important properties such as phonon spectrum, group velocity, and density of states can be computed from the time history of the atomic trajectory and its velocity [320]. In principle, such information can be incorporated into the Boltzmann equation to treat heat transfer mesoscale systems that cannot be simulated directly by the MD methods. In addition, MD simulation can be the most ideal tool to study the imperfect interface and anharmonic scattering while lattice dynamics or acoustic modeling can only model the process qualitatively by assuming idealized interfaces and neglecting all anharmonic scattering in nanostructures. Some MD studies have examined thermal boundary resistance [321–324].

One theoretical drawback of the classical MD method is that it does not include the quantum statistics and quantum scattering mechanism. Such a drawback limits the MD results to high temperatures when the quantum effects are negligible. There have also been theoretical attempts to correct the difference between quantum and classical simulations. The first correction is on the temperature. Other modeling suggests that quantum effects on scattering are not large. This topic, however, needs further study.

5.3.3. Fluctuation–Dissipation Theorem Based Wave Approach

A well-established method to treat both the absorption, emission, and wave effects in thermal radiation is the fluctuaction–dissipation theorem based approach established by Rytov [325]. This approach was used to formulate the problem of radiation heat transfer between two closely spaced surfaces [14], radiative transfer in absorbing and emitting thin films [326], and recently was used for treating radiation exchange between surfaces due to surface modes [147, 151]. In this approach, the emission is related to the fluctuation of microscopic current sources, the strength of which is related to the local temperature through the fluctuation–dissipation theorem. The propagation of the electromagnetic field generated by each current source is solved based on the Maxwell equations. This approach, in theory, can be extended to acoustic waves and maybe even electron waves, but no such attempts have been reported.

5.4. Multicarrier and Multidimensional Transport

Heat transfer and energy conversion in electronic and thermoelectric devices often involve multicarriers and multiple length scales. For example, heat is typically generated in a nanometer-scale region at the drain of a MOSFET when hot electrons relax their energy to phonons and is subsequently conducted through the millimeter thick substrate to the surroundings. In this section, we will discuss the modeling of nonequilibrium electron and phonon transport and the multidimensional transport problems.

5.4.1. Coupled Electron–Phonon Transport

Traditionally, it is assumed that electrons and phonons are under local equilibrium in modeling transport phenomena in a solid. However, the equilibrium between electrons and phonons can be disrupted in many cases. For example, in the presence of a sufficiently high electric field, electrons can be energized and thrown far out of equilibrium from phonons [127]. This is termed the hot electron effect. Such a nonequilibrium condition becomes important for microelectronic devices because the electric fields become higher as the feature size shrinks. In the case of laser–material interactions, the electrons can be thrown out of equilibrium from lattice due to excitation by ultrashort laser pulses [126, 327].

For coupled electron and phonon transport in nanostructures, the electron and phonon Boltzmann equations should be solved simultaneously. This approach, however, has not been taken due to the numerical complexity involved. In the past, the electron hydrodynamic equations are often used for modeling electron transport for practical device simulation, coupled to the diffusion treatment for the phonon heat conduction [328–331]. The hydrodynamic equations assume that electrons and phonons are in local equilibrium with their own pools. In a quite similar way, Qiu and Tien studied conversion and transport phenomena during ultrafast processes such as femtosecond to nanosecond laser–material interactions based on treating electrons and phonons as separate systems [126, 332]. Their theoretical and experimental studies show that it is the nonequilibrium nature between electrons and phonons during the femtosecond laser–metal interactions, rather than the hyperbolic heat conduction effect, that dominates the observed experimental phenomena. However, in these models, the ballistic transports are not counted, neither for electrons nor for phonons. Sometimes the electron transport is solved from the full-blown Boltzmann equation or Monte–Carlo simulation to count the ballistic nature of electrons [333]. However, very little work has been done to study the effect of ballistic phonon transport on electron transport. Concurrent modeling of the nonequilibrium electron and phonon transport based Boltzmann equation is desirable.

Along the same line, the coupled electron–phonon transport is the basis for solid-state energy conversion such as thermoelectric and thermionic cooling and power generation. Taking thermoelectric cooling as an example, electrons take energy away from phonons at a metal–semiconductor interface, carrying it to the hot side, and rejecting it to phonons. Although the macroscopic thermoelectric phenomenon is well understood [334], some questions remain, such as what happens at the interface and how phonons are cooled and heated by electron systems. There are a few papers dealing with nonequilibrium electron and phonon transport in thermoelectric research [46, 335, 336] but the full picture of coupled electron–phonon transport near the interface has yet to be clarified.

5.4.2. Multilength Scale and Multidimensional Transport

Heat conduction in most devices is multidimensional and involves length scales from nanoscale to macroscale. Modeling the energy transport and conversion processes in such multiple length scale devices is very challenging. The solution of the Boltzmann equation has been limited to a few a simple geometrical configurations such as thin films and superlattices [76, 337]. In principle, the discrete ordinates method, the spherical harmonics method, and the Monte Carlo method developed for thermal radiation

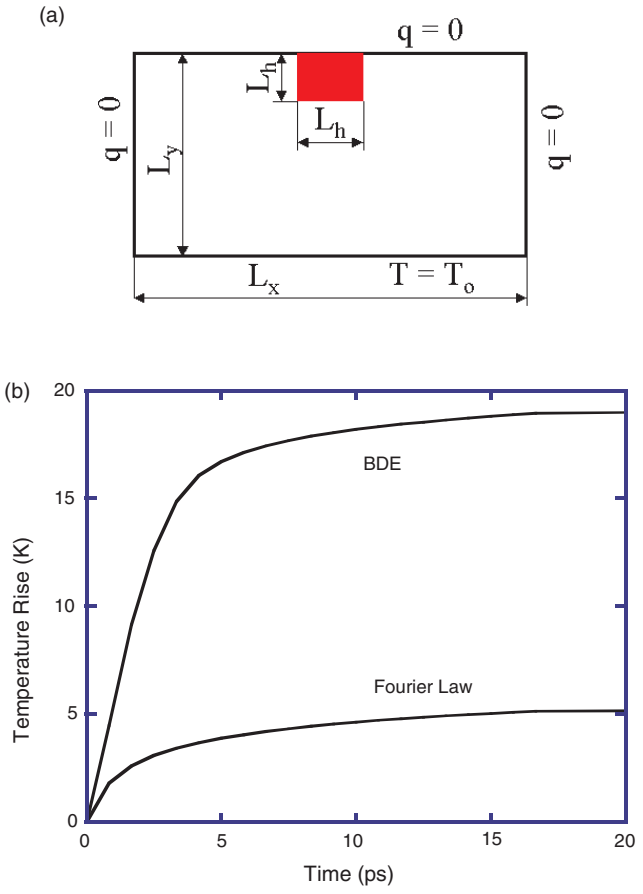


Figure 18. Comparison of the peak temperature rise of nanoscale heat source. (a) a 10 nm × 10 nm ×1μm hot strip embedded in the silicon substrate and (b) the peak temperature predicted by the Fourier law is much smaller than that predicted by the ballistic–diffusive equations. Reprinted with permission from [121], R. G. Yang, et al., in "Proceedings of the 12th International Heat Transfer Conference" (J. Taine, Ed.), Vol. 1, pp. 579–584. Elsevier, Paris, 2002. © 2002, Elsevier.

[253] can be used to solve the Boltzmann transport equation. However, fewer studies have gone beyond nonplanar and multidimensional geometries [117, 120, 121, 271, 283, 338, 339].

Direct numerical solution of the Boltzmann equation, however, is usually slow. Approximate methods that are capable of capturing the major size effects but easier to implement are thus desirable. One such method, for example, is to focus on the interface region only by introducing appropriate boundary conditions, while away from the boundaries the usual diffusion or drift-diffusion equations are used. Examples are the velocity slip boundary condition for rarefied gas flow [256] and the Deissler temperature jump boundary condition for photon radiation heat transfer [25]. Similar interface conditions, called diffusion–transmission interface conditions, have been introduced for phonon and electron transport [340].

Another approximation that provides further improvements compared to the introduction of special interface conditions is to split the carriers inside the medium into two groups—a ballistic component and a diffusive component. The ballistic component originating from boundaries and nanoscale heat sources experiences outscattering only. This group can be explicitly solved through a path integral. The transport of the scattered and excited heat carriers is treated as a diffusive component with the diffusion approximation. This approach was first used for thermal radiation [253] and was recently employed by Chen to develop the ballistic–diffusive heat conduction equations. Yang et al. [121, 338] applied the ballistic–diffusive equations to study the phonon rarefaction effect, the heat transfer surrounding a nanoscale heat source. The heat source is similar to the phonon hot spot generated in a MOSFET device. A 10 nm $\times$ 10 nm $\times$ 1μm hot strip is embedded in the silicon substrate (Fig. 18a). The power generation rate is 1×10^{19} W/m^3, typically for a period of $\sim$10 ps. Figure 18b shows the highest temperature rise in the device as a function of time. Both Fourier law and ballistic–diffusive equations predict a saturation temperature rise after 10 ps. The peak temperature predicted by the Fourier law is much smaller than that predicted by the ballistic–diffusive equation, due to the rarefied phonon heat conduction effect discussed in Section 3. The experiment by Sverdrup and co-workers [123] also demonstrates such a phonon rarefaction effect.

6. SUMMARY

In this chapter, we discussed the uniqueness of heat transfer phenomena in nanostructures. Figure 6 summarizes different heat transfer regimes that one may encounter at nanoscale. When the structure characteristic lengths are comparable to the mean free path, size effects become important. In the classical size effect regime, the phase of the energy carriers can be neglected while it must be included in the quantum size effect regime. The distinction between the classical and the quantum size effect regime depends on the phase-breaking strength of the interface scattering processes, which is not well understood now. Thermal wavelength provides another length scale to judge whether the bulk energy spectrum is a good approximation. This length scale, however, must be used carefully because although the average energy spectrum does not change much when the structure is much larger than the thermal wavelength, the group velocity in specific transport directions can have significant changes, as in the case of superlattices. Some examples of nanoscale heat transfer phenomena in the quantum regime include interference, tunneling, and even coherent thermal emission. In the classical size effect regime, diffuse interface scattering tends to reduce the heat transfer through the system.

Nanoscale heat transfer phenomena can have significant implications for a variety of contemporary technologies, some undesirable as the classical size effects on the thermal management of microelectronics and optoelectronic devices, while others can be utilized to improve the data storage density or energy conversion efficiency, as in thermoelectric cooling and power generation and in thermophotovoltaic energy conversion. Nanoscale heat transfer is also related to nanomaterial synthesis and nanofabrication and some emerging biotechnologies.

Experimental and modeling tools for nanoscale heat transfer are discussed. Various measurement techniques have been developed in the past for the thermophysical property determination of thin films and superlattices. Experimental techniques for characterizing nanowires and other nanostructures, such as nanowires and nanoparticles, still need further development. Impressive progresss has been made in thermal imaging at nanoscale based on the scanning probe microscope but the technique has stayed mainly at a qualitative level. With regard to modeling, the classical size effects region and purely ballistic transport regimes can be treated reasonably well. Key challenges exist for the regime when both wave effects and scattering coexist because scattering can change the energy spectrum of nanostructures. At room temperature, experimental studies so far suggest that phonon transport falls mostly into the classical size effect regime while electron transport can fall into the quantized incoherent transport regime. From the discussions, it is clear that much work remains to be done in both experimental and modeling arenas.

GLOSSARY

Acoustic phonons The lower branches of phonon modes in the dispersion curves.

Blackbody radiation The radiation emitted by a blackbody. A blackbody is a virtual object that absorbs all radiation incidents on it.

Dispersion relation Relates the frequency of a carrier, ω_k, to its wavevector k. The *group velocity* of the carrier can be found from the dispersion relation through $v_k = \frac{\partial \omega_k}{\partial k}$.

Elastic scattering There is no net change of energy of each involved particle before and after an elastic scattering process.

Inelastic scattering The energy of some of participating carriers changes before and after an inelastic scattering process. Refer to elastic scattering.

Mean free path The average distance a particle moves between succesive collisions.

Optical phonons The upper branches of phonon modes in the dispersion curves, refer to acoustic phonons.

Phonons Quanta of lattice waves in solids.

Phonon-polariton The hybrid mode of the coupled phonon and electromagnetic transverse wave field.

Phonon rarefaction Occurs when the heating/cooling region is much smaller than the phonon mean free path.

Stokes/Anti-Stokes shift Represents a change in energy (frequency) of photon scattered by a substance due to photon-phonon interaction. The scattered photon may have a smaller frequency (Stokes shift) or higher frequency (anti-Stokes shift) depending whether a phonon has been generated or absorbed during the scattering process.

Thermal boundary resistance or Kapitza resistance Used to describe the additional thermal resistance due to the reflection of energy carriers at an interface when the heat transports across the interface.

ACKNOWLEDGMENTS

The authors are grateful to M. S. Dresselhaus and members of the Nanoengineering Group at MIT Rohsenow Heat Transfer Lab for stimulating discussions. G.C. gratefully acknowledges the support of NSF (CTS-0129088), DOE (DE-FG-02ER45977), ONR MURI on Thermoelectrics, and ONR MURI on Electromagnetic Metamaterials through UCLA, and by Draper Lab.

REFERENCES

1. C.-L. Tien, A. Majumdar, and F. M. Gerner, "Microscale Energy Transport." Taylor & Francis, Washington, 1998.
2. G. Chen, "Nanoscale Heat Transfer and Energy Conversion." Oxford Univ. Press, in press.
3. K. E. Goodson and Y. S. Ju, *Annu. Rev. Mater. Sci.* 29, 261 (1999).
4. G. Chen, *Annu. Rev. Heat Transf.* 7, 69 (1996).
5. T. Tritt, "Recent Trends in Thermoelectric Materials Research," Semiconductors and Semimetals Series, Vol. 69–71. Academic Press, San Diego, 2001.
6. G. Chen and A. Shakouri, *Trans. ASME, J. Heat Transf.* 124, 242 (2002).
7. R. S. DiMatteo, P. Greiff, S. L. Finberg, K. Young-Waithe, H. K. H. Choy, M. M. Masaki, and C. G. Fonstad, *Appl. Phys. Lett.* **79**, 1894 (2001).
8. G. Binning, M. Despont, U. Drechsler, W. Häberle, M. Lutwyche, P. Vettiger, H. J. Mamin, B. W. Choi, and T. W. Kenny, *Appl. Phys. Lett.* 74, 1329 (1999).
9. M. I. Flik, B. I. Choi, and K. E. Goodson, *Trans. ASME, J. Heat Transf.* 114, 667 (1992).
10. C. L. Tien and G. Chen, *Trans. ASME, J. Heat Transf.* **116**, 799 (1994).
11. G. Chen, "Semiconductors and Semimetals," Vol. 71, p. 203. Academic Press, San Diego, 2001.
12. D. G. Cahill, W. K. Ford, K. E. Goodson, G. D. Mahan, A. Majumdar, H. J. Maris, and R. Merlin, *J. Appl. Phys.* 93, 743 (2003).
13. G. A. Domoto, R. F. Boehm, and C. L. Tien, *Trans. ASME, J. Heat Transf.* 92, 412 (1972).
14. D. Polder and M. Von Hove, *Phys. Rev.* 4, 3303 (1971).
15. F. P. Incropera and D. P. DeWitt, "Fundamentals of Heat and Mass Transfer," 5th ed. Wiley, New York, 2002.
16. C. Kittel, "Introduction to Solid State Physics," 7th ed. Wiley, New York, 1996.
17. N. W. Aschroft, and N. D. Mermin, "Solid State Physics." Sauders, Forth Worth, PA, 1976.
18. R. Berman, "Thermal Conduction in Solids." Clarendon, Oxford, 1976.
19. J. M. Ziman, "Electrons and Phonons." Clarendon, Oxford, 1960.
20. P. Hyldgaard and G. D. Mahan, "Thermal Conductivity." Technomic, Lancaster, UK, 1996.
21. G. Chen, *Trans. ASME, J. Heat Transf.* 119, 220 (1997).
22. Y. S. Ju and K. E. Goodson, *Appl. Phys. Lett.* 74, 3005 (1999).
23. D. K. Ferry and S. M. Goodnick, "Transport in Nanostructures." Cambridge Univ. Press, Cambridge, UK, 1999.
24. M. Born and E. Wolf, "Principles of Optics," 6th ed. Pergamon Press, Oxford, 1980.
25. R. Siegel and J. R. Howell, "Thermal Radiation Heat Transfer." Hemisphere, Washington, 1992.
26. C. L. Mehta, *Nuovo Cimento* 21, 401 (1963).
27. G. Chen and C. L. Tien, *Trans. ASME, J. Heat Transf.* 114, 636 (1992).
28. G. Chen, *Trans. ASME, J. Heat Transf.* 121, 945 (1999).
29. S. Schaaf, and P. Chambre, "Flow of Rarefied Gases." Princeton Univ. Press, Princeton, NJ, 1961.
30. L. D. Hicks and M. S. Dresselhaus, *Phys. Rev. B* 47, 16631 (1993).
31. M. S. Dresselhaus, Y. M. Lin, S. B. Cronin, O. Rabin, M. R. Black, G. Dresselhaus, and T. Koga, "Semiconductors and Semimetals," Vol. 71, p. 1. Academic Press, San Diego, 2001.
32. B. Yang and G. Chen, *Microscale Thermophys. Eng.* 5, 107 (2001).
33. H. Raether, "Surface Plasmons." Springer-Verlag, Berlin, 1998.
34. D. L. Mills and E. Burnstein, *Rep. Progr. Phys.* 37, 817 (1974).
35. C. K. Campbell, "Surface Acoustic Wave Devices for Mobile and Wireless Communications." Academic Press, San Diego, 1998.
36. J. Grefett, R. Carminati, K. Joulain, J.-P. Mulet, S. Malnguy, and Y. Chen, *Nature* 416, 61 (2002).
37. R. Hillenbrand, T. Taubner, and F. Keilmann, *Nature* 418, 159 (2002)
38. L. Esaki and R. Tsu, *IBM J. Res. Dev.* 14, 61 (1970).
39. C. Weisbuch and B. Vinter, "Quantum Semiconductor Structures: Fundamentals and Applications." Academic Press, Boston, 1991.
40. E. Yablonovitch, *Phys. Rev. Lett.* 58, 2059 (1986).
41. J. Joannapoulos, R. D. Meade, and J. N. Winn, "Photonic Crystals." Princeton Univ. Press, Princeton, NJ 1995.
42. E. T. Swartz and R. O. Pohl, *Rev. Mod. Phys.* 61, 605 (1989).
43. P. L. Kapitza, *Zh. Eksp. Teor. Fiz.* 11, 1 (1941) [*J. Phys. (USSR)* 4, 181 (1941)]; in "Collected Papers of P. L. Kapitza" (D. der Haar, Ed.), Vol. 2, p. 581. Pergamon, Oxford, 1965.
44. W. A. Little, *Can. J. Phys.* 37, 334 (1959).
45. C. Dames and G. Chen, *J. Appl. Phys.* 95, 682 (2004).
46. G. Chen and T. Zeng, *Microscale Thermophys. Eng.* 5, 71 (2001).
47. J. A. Katerberg, C. L. Reynolds, Jr., and A. C. Anderson, *Phys. Rev. B* 16, 673 (1977).
48. G. Chen, *Phys. Rev. B* 57, 14958 (1998).
49. J. Freund and G. Chen, unpublished.
50. S. M. Lee, D. G. Cahill, and R. Venkatasubramanian, *Appl. Phys. Lett.* 70, 2957 (1997).
51. B. Yang and G. Chen, in "Chemistry, Physics, and Materials Science for Thermoelectric Materials: Beyond Bismuth Telluride" (M. G. Kanatzidis, T. P. Hogan, and S. D. Mahanti, Eds.). Kluwer Academic/Plenum, New York, 2003.
52. G. D. Mahan and M. Bartkowiak, *Appl. Phys. Lett.* 74, 953 (1999).
53. K. H. Yoo and A. C. Anderson, *J. Low Temp. Phys.* 63, 269 (1986).
54. K. Fuchs, *Proc. Cambridge Philos. Soc.* 34, 100 (1938).
55. E. H. Sondheimer, *Phys. Rev.* 80, 401 (1950).

56. C. R. Tellier and A. J. Tosser, "Size Effects in Thin Films." Elsevier, New York, 1982.
57. B. Yang and G. Chen, *Phys. Low-Dimens. Struct.* 5/6, 37 (2000).
58. A. Balandin and K. L. Wang, *Phys. Rev. B.* 58, 1544 (1998).
59. S. Volz and D. Lemonnier, *Phys. Low-Dimens. Struct.* 5/6, 91 (2000).
60. M. Asheghi, M. N. Touzelbaev, K. E. Goodson, Y. K. Leung, and S. S. Wong, *Trans. ASME, J. Heat Transf.* 120, 31 (1998).
61. M. Asheghi, K. Kurabayashi, R. Kasnavi, and K. E. Goodson, *J. Appl. Phys.* 91, 5079 (2002).
62. D. G. Song and G. Chen, in "Proceedings of ICT 2002: 21st International Conference on Thermoelectrics." p. 292, IEEE, New York, 2002.
63. F. Volklein and H. Baltes, *J. Microelectromech. Syst.* 1, 193 (1992).
64. F. Volklein and T. Starz, in "XVI ICT '97: Proceedings of ICT'97, the 16th International Conference on Thermoelectrics," p. 711. IEEE, New York, 1998.
65. Y. C. Tai, C. H. Mastrangelo, and R. S. Muller, *J. Appl. Phys.* 63, 1442 (1988).
66. K. E. Goodson, O. W. Kading, M. Rosner, and R. Zachai, *Appl. Phys. Lett.* 66, 3134 (1995).
67. S. Uma, A. D. McConnell, M. Asheghi, K. Kurabayashi, and K. E. Goodson, *Intl. J. Thermophys.* 22, 605 (2001).
68. A. D. McConnell, S. Uma, and K. E. Goodson, *J. Microelectromech. Syst.* 10, 360 (2001).
69. M. Werner, T. Kohler, S. Mietke, E. Worner, C. Johnston, and H.-J. Fecht, *Proc. SPIE* 4703, 199 (2002).
70. B. C. Daly, H. J. Maris, A. V. Nurmikko, M. Kuball, and J. Han, *J. Appl. Phys.* 92, 3820 (2002).
71. V. Baier, F. Volklein, *Phys. Status Solidi A* 118, K69 (1990).
72. M. Rohde, *Thin Solid Films* 238, 199 (1994).
73. E. J. Gonzalez, J. E. Bonevich, G. R. Stafford, G. White, and D. Josell, *J. Mater. Res.* 15, 764 (2000).
74. S.-M. Lee and D. G. Cahill, *J. Appl. Phys.* 81, 2590 (1997).
75. S.-M. Lee and D. G. Cahill, *Phys. Rev. B* 52, 253 (1995).
76. Y. S. Ju and K. E. Goodson, *J. Appl. Phys.* 85, 7130 (1999).
77. M. B. Kleiner, S. A. Kuhn, and W. Weber, *IEEE Trans. Electron Devices* 43, 1602 (1996).
78. A. J. Griffin, Jr., F. R. Brotzen, and P. L. Loos, *J. Appl. Phys.* 75, 3761 (1994).
79. T. Yao, *Appl. Phys. Lett.* 51, 1298 (1987).
80. W. S. Capinski, H. J. Maris, T. Ruf, M. Cardona, K. Ploog, and D. S. Katzer, *Phys. Rev. B* 59, 8105 (1999).
81. T. Borca-Tasciuc, D. W. Song, J. R. Meyer, I. Vurgaftman, M.-J. Yang, B. Z. Nosho, L. J. Whitman, H. Lee, R. U. Martinelli, G. W. Turner, M. J. Manfra, and G. Chen, *J. Appl. Phys.* 92, 4994 (2002).
82. B. Yang, J. L. Liu, K. L. Wang, and G. Chen, *Appl. Phys. Lett.* 80, 1758 (2002).
83. W. L. Liu, T. Borca-Tasciuc, G. Chen, J. L. Liu, and K. L. Wang, *J. Nanosci. Nanotechnol.* 1, 39 (2001).
84. R. Venkatasubramanian, *Phys. Rev. B* 61, 3091 (2000).
85. T. Borca-Tasciuc, D. Song, J. L. Liu, G. Chen, K. L. Wang, X. Sun, M. S. Dresselhaus, T. Radetic, and R. Gronsky, in "Thermoelectric Materials 1998—Next Generation Materials for Small-Scale Refrigeration and Power Generation Applications. Symposium Proceedings" (T. Tritt, M. G. Kanatzidis, G. D. Mahan, H. B. Lion, Jr., Eds.), Vol. 524, p. 473. Material Research Society, Warrendale, PA, 1999.
86. V. Narayanamurti, *J. Phys. Colloq.* 45, C5/157 (1984).
87. G. Scamarcio, V. Spagnolo, E. Molinari, L. Tapfer, L. Sorba, G. Bratina, and A. Franciosi, *Phys. Rev. B* 46, 7296 (1992).
88. G. Fasol, D. Richards, J. D. White, K. Ploog, C. J. Gibbings, and C. G. Tuppen, *Semicond. Sci. Technol.* 5, 1168 (1990).
89. P. Hyldegaard and G. D. Mahan, *Phys. Rev. B* 56, 10754 (1997).
90. S. Tamura and J. P. Wolfe, *Phys. Rev. B* 38, 5610 (1998).
91. W. E. Bies, R. J. Radtke, and H. Ehrenreich, *J. Appl. Phys.* 88, 1498 (2000).
92. S. Tamura, Y. Tanaka, and H. J. Maris, *Phys. Rev. B* 60, 2627 (1999).
93. G. Chen and M. Neagu, *Appl. Phys. Lett.* 71, 2761 (1997).
94. J. Seyler and M. N. Wybourne, *Phys. Rev. Lett.* 69, 1427 (1992).
95. A. Potts, M. J. Kelly, C. G. Smith, D. G. Hasko, J. R. A. Cleaver, H. Ahmed, D. C. Peacock, J. E. F. Frost, D. A. Ritchie, and G. A. C. Jones, *J. Phys. Condens. Matter* 2, 1817 (1990).
96. A. Potts, M. J. Kelly, D. G. Hasko, J. R. A. Cleaver, H. Ahmed, D. C. Peacock, J. E. F. Frost, D. A. Ritchie, and G. A. C. Jones, *Superlatt. Microstruct.* 9, 315 (1991).
97. A. Potts, M. J. Kelly, D. G. Hasko, J. R. A. Cleaver, H. Ahmed, J. E. F. Frost, D. A. Ritchie, and G. A. C. Jones, *Semicond. Sci. Technol.* 7, B231 (1992).
98. J. Zou and A. Balandin, *J. Appl. Phys.* 89, 2932 (2001).
99. P. Kim, L. Shi, A. Majumdar, and P. L. McEuen, *Phys. Rev. Lett.* 87, 215502 (2001).
100. T. S. Tighe, J. M. Worlock, and M. L. Roukes, *Appl. Phys. Lett.* 70, 2689 (1997).
101. K. Schwab, E. A. Henriksen, J. M. Worlock, and M. L. Roukes, *Nature* 404, 974 (2001).
102. D. E. Angelescu, M. C. Cross, and M. L. Roukes, *Superlatt. Microstruct.* 23, 673 (1998).
103. X. Lü, W. Z. Shen, and J. H. Chu, *J. Appl. Phys.* 91, 1542 (2002).
104. S. Huxtable, D. Li, A. Abramson, A. Miner, W. Kim, M. Chapp, and A. Majumdar, presented at the "US–Japan Nanotherm Seminar: Nanoscale Thermal Science and Engineering," 2002.
105. S. G. Volz and G. Chen, *Appl. Phys. Lett.* 75, 2056 (1999).
106. S. G. Walkauskas, D. A. Broido, K. Kempa, and T. L. Reinecke, *J. Appl. Phys.* 85, 2579 (1999).
107. M. S. Dresselhaus, G. Dresselhaus, and P. C. Eklund, "Science of Fullerenes and Carbon Nanotubes." Academic Press, San Diego, 1996.
108. R. S. Ruoff and D. C. Lorents, *Carbon* 33, 925 (1995).
109. S. Berber, Y.-K. Kwon, and D. Tomanek, *Phys. Rev. Lett.* 84, 4631 (2000).
110. C. Jianwei, T. Cagin, and W. A. Goddard III, *Nanotechnology* 11, 65 (2000).
111. M. A. Osman and D. Srivastava, *Nanotechnology* 12, 21 (2001).
112. J. F. Moreland, J. B. Freund, and G. Chen, *Microscale Thermophys. Eng.*, in press.
113. S. Marayama, *Microscale Thermophys. Eng.* 7, 41 (2003).
114. J. Hone, M. Whitney, C. Piskoti, and A. Zettl, *Phys. Rev. B* 59, R2514 (1999).
115. T. Borca-Tasciuc, C. L. Hapenciuc, B. Wei, R. Vajtai, and P. M. Ajayan, in "Proceedings of the ASME Heat Transfer Division," 2002 (CD-ROM).
116. D. J. Yang, Q. Zhang, G. Chen, S. F. Yoon, J. Ahn, S. G. Wang, Q. Zhou, Q. Wang, and J. Q. Li, *Phys. Rev. B* 66, 165440 (2002).
117. G. Chen, *Trans. ASME, J. Heat Transf.* 118, 539 (1996).
118. G. D. Mahan and F. Claro, *Phys. Rev. B* 38, 1963 (1988).
119. F. Claro and G. D. Mahan, *J. Appl. Phys.* 66, 4213 (1989).
120. P. G. Sverdrup, Y. S. Ju, and K. E. Goodson, *Trans. ASME, J. Heat Transf.* 123, 130 (2001).
121. R. G. Yang, G. Chen, Y. Taur, in "Proceeding of the 12th International Heat Transfer Conference" (J. Taine, Ed.), Vol. 1, pp. 579–584. Elsevier, New York, 2002.
122. Y. Z. Li, L. Vazquez, R. Piner, R. P. Andres, and R. Reifenberger, *Appl. Phys. Lett.* 54, 1424 (1989).
123. P. G. Sverdrup, S. Sinha, S. Uma, M. Asheghi, and K. E. Goodson, *Appl. Phys. Lett.* 78, 3331 (2001).
124. M. I. Kaganov, I. M. Lifshitz, and L. V. Tanatarov, *Zh. Eksp. Teor. Fiz.* 31, 232 (1956) [*Sov. Phys. JETP* 4, 173 (1957)].
125. R. W. Schoenlein, W. Z. Lin, J. G. Fujimoto, and G. L. Easley, *Phys. Rev. Lett.* 58, 1680 (1987).

126. T. Q. Qiu and C. L. Tien, *Trans. ASME, J. Heat Transf.* 115, 835 (1993).
127. K. Banerjee, A. Amerasekera, G. Dixit, N. Cheung, and C. Hu, "International Electron Devices Meeting, Technical Digest," p. 216. IEEE, Piscataway, NJ, 1997.
128. A. Majumdar, K. Fushinobu, and K. Hijikata, *J. Appl. Phys.* 77, 6686 (1995).
129. E. Pop, K. Banerjee, P. Sverdrup, R. Dutton, and K. Goodson, in "International Electron Devices Meeting, Technical Digest," Vol. 951, p. 31.1.1. IEEE, Piscataway, NJ, 2001.
130. G. D. Mahan, *Semicond. Semimetals* 69, p. 172.
131. G. Chen and C. L. Tien, *Trans. ASME, J. Heat Transf.* 114, 636 (1992).
132. K. Richter, G. Chen, and C. L. Tien, *Opt. Eng.* 32, 1897 (1993).
133. P. Y. Wong, C. K. Hess, and I. N. Miaoulis, *Opt. Eng.* 34, 1776 (1995).
134. Z. M. Zhang, B. I. Choi, T. A. Le, M. I. Flik, M. P. Siegal, and J. M. Phillips, *Trans. ASME, J. Heat Transf.* 114, 644 (1992).
135. Z. M. Zhang and M. I. Flik, *IEEE Trans. Appl. Supercond.* 3, 1604 (1993).
136. A. R. Kumar, Z. M. Zhang, V. A. Boychev, D. B. Tanner, L. R. Vale, and D. A. Rudman, *Trans. ASME, J. Heat Transf.* 12, 844 (1999).
137. A. R. Kumar, V. A. Boychev, Z. M. Zhang, and D. B. Tanner, *Trans. ASME, J. Heat Transf.* 122, 785 (2000).
138. C. F. Bohren and D. R. Huffman, "Absorption and Scattering of Light by Small Particles." Wiley, New York, 1983.
139. S. Y. Lin, J. G. Fleming, D. L. Hetherington, B. K. Smith, R. Biswas, K. M. Ho, M. M. Sigalas, W. Zubrzycki, S. R. Kurtz, and J. Bur, *Nature* 394, 251 (1998).
140. G. Subramania, R. Biswas, K. Constant, M. M. Sigalas, K. M. Ho, *Phys. Rev. B* 63, 235111 (2001).
141. A. Rosenberg, R. J. Tonucci, and E. L. Shirley, *J. Appl. Phys.* 82, 6354 (1997).
142. S. D. Hart, G. R. Maskaly, B. Temelkuran, P. H. Prideaux, J. D. Joannopoulos, and Y. Fink, *Science* 296, 510 (2002).
143. C. L. Tien and G. R. Cunnington, *Adv. Heat Transf.* 9, 349 (1973).
144. E. G. Cravolho, C. L. Tien, and R. P. Caren, *Trans. ASME, J. Heat Transf.* 89, 351 (1967).
145. J. B. Pendry, *J. Phys. Condens. Matter* 11, 6621 (1999).
146. C. M. Hargreaves, *Phys. Lett.* 30A, 491 (1969).
147. R. Carminati and J. J. Greffet, *Phys. Rev. Lett.* 82, 1660 (1999).
148. A. V. Shchegrov, K. Joulain, R. Carminati, and J. J. Greffet, *Phys. Rev. Lett.* 85, 1548 (2000).
149. J. Le-Gall, M. Olivier, and J. J. Greffet, *Phys. Rev. B* 55, 10105 (1997).
150. J. P. Mulet, K. Joulain, R. Carminati, and J. J. Greffet, *Appl. Phys. Lett.* 78, 2931 (2001).
151. A. Narayamaswany and G. Chen, *Appl. Phys. Lett.* 82, 3544 (2003).
152. G. A. Domoto and C. L. Tien, *Trans. ASME, J. Heat Transf.* 92, 399 (1970).
153. E. H. Kennard, "Kinetic Theory of Gases." McGraw–Hill, New York, 1938.
154. J. Huisken and E. H. K. Stelzer, *Opt. Lett.* 27, 1223 (2002).
155. J. N. Israelachvili, "Intermolecular and Surface Forces." Academic Press, San Diego/London, 1992.
156. G. E. Karniadakis and A. Beskok, "Micro Flows, Fundamentals and Simmulations." Springer-Verlag, New York, 2001.
157. C.-M. Ho and Y.-C. Tai, *Annu. Rev. Fluid Mech.* 30, 579 (1998).
158. J. C. Maxwell, "A Treatise on Electricity and Magnetism," 2nd ed. Clarendon, Oxford, 1881.
159. J. A. Eastman, S. U. S. Choi, S. Li, W. Yu, and L. J. Thompson, *Appl. Phys. Lett.* 78, 718 (2001).
160. P. Keblinski, S. R. Phillpot, S. U. Choi, and J. A. Eastman, *Int. J. Heat Mass Tranf.* 45, 855 (2002).
161. Y. Taur, C. H. Wann, and D. J. Frank, in "International Electron Devices Meeting 1998, Technical Digest," Vol. 1080, p. 789. IEEE, Piscataway, NJ, 1998.
162. G. Hasnain, K. Tai, L. Yang, Y. H. Wang, R. J. Fischer, J. D. Wynn, B. Weir, N. K. Dutta, and A. Y. Cho, *IEEE J. Quantum Electron.* 27, 1377 (1991).
163. B. Tell, K. F. Brown-Goebeler, and R. E. Leibenguth, *IEEE Photonics Technol. Lett.* 4, 521 (1992).
164. G. Chen, *J. Appl. Phys.* 77, 4152 (1995).
165. H. D. Summers, J. Wu, and J. S. Roberts, *IEE Proc. Optoelectron.* 148, 261 (2001).
166. G. Chen and C. L. Tien, *J. Appl. Phys.* 74, 2167 (1993).
167. J. Tatum, in "2001 IEEE Emerging Technologies Symposium on Broadband Communications for the Internet Era, Symposium Digest" (P. Winson, Ed.), Vol. 161, p. 58. IEEE, Piscataway, NJ, 2001.
168. S. Fukui, and R. Kaneko, *IEEE Trans. Magn.* 24, 2751 (1988).
169. S. Fukui, K. Yamane, and H. Masuda, *IEEE Trans. Magn.* 37, 1845 (2001).
170. E. Grochowski and R. E. Fontana, Jr., in "Seventh Biennial IEEE International Nonvolatile Memory Technology Conference Proceedings," p. 8. IEEE, New York, 1998.
171. S. Sun, C. B. Murray, D. Weller, L. Folks, and A. Moser, *Science* 287, 1989 (2000).
172. W. P. King, T. W. Kenny, K. E. Goodson, G. Cross, M. Despont, U. Durig, H. Rothuizen, G. K. Binnig, and P. Vettiger, *Appl. Phys. Lett.* 78, 1300 (2001).
173. U. Duriga, G. Cross, M. Despont, U. Drechsler, W. Haberle, M. I. Lutwyche, H. Rothuizen, R. Stutz, R. Widmer, P. Vettiger, G. K. Binnig, W. P. King, and K. E. Goodson, *Tribol. Lett.* 9, 25 (2000).
174. A. B. Marchant, "Optical Recording." Addison–Wesley, Reading, MA, 1972.
175. M. Terao, Y. Miyauchi, K. Andoo, H. Yasuoka, and R. Tamura, *Proc. SPIE* 1078, 2 (1989).
176. R. Ito, Y. Tsunoda, H. Ohta, S. Kubota, K. Ogawa, M. Okuda, M. Irie, S. Mitsumori, and H. Nshihara, *Jpn. J. Appl. Phys.* 32, 5185 (1993).
177. H. J. Goldsmid, "Applications of Thermoelectricity." Methuen/Wiley, London/New York, 1960.
178. L. D. Hicks and M. S. Dresselhaus, *Phys. Rev. B* 47, 12727 (1993).
179. T. Koga, X. Sun, S. B. Cronin, M. S. Dresselhaus, K. L. Wang, and G. Chen, *J. Comput.-Aided Mater. Des.* 82, 830 (1997).
180. T. Koga, X. Sun, S. B. Cronin, and M. S. Dresselhaus, *Appl. Phys. Lett.* 75, 2438 (1999).
181. X. Sun, Z. Zhang, and M. S. Dresselhaus, *Appl. Phys. Lett.* 74, 4005 (1999).
182. T. C. Harman, P. J. Taylor, D. L. Spears, and M. P. Walsh, *J. Electron. Mater.* 29, L1 (2000).
183. M. S. Dresselhaus, G. D. Dresselhaus, X. Sun, Z. Zhang, S. B. Cronin, T. Koga, J. Y. Ying, and G. Chen, *Microscale Thermophys. Eng.* 3, 89 (1999).
184. G. Chen, M. S. Dresselhaus, J.-P. Fleurial, and T. Caillat, *Intl. Mat. Rev.* 48, 45 (2003).
185. R. Venkatasubramanian, E. Siivola, T. Colpitts, and B. O'Quinn, *Nature* 413, 597 (2001).
186. T. C. Harman, P. J. Taylor, M. P. Walsh, and B. E. LaForge, *Science* 297, 2229 (2002).
187. J. G. Fleming, S. Y. Lin, I. El-Kady, R. Biswas, and K. M. Ho, *Nature* 417, 52 (2002).
188. M. D. Whale and E. G. Cravalho, *IEEE Trans. Energy Convers.* 17, 130 (2002).
189. M. Kreiter, J. Oster, R. Sambles, S. Herminghaus, S. Mittler-Neher, and W. Knoll, *Opt. Commun.* 168, 117 (1999).
190. B. A. Ridley, B. Nivi, and J. M. Jacobson, *Science* 286, 746 (1999).

191. S. Hector and P. Mangat, *J. Vac. Sci. Technol. B* 19, 2612 (2001).
192. C. Grigoropoulous, in "Proc. DOE 20th Sympos. Energy Eng. Sci.," Argonne, IL, 2002, p. 154.
193. S. Y. Chou, C. Keimel, and J. Gu, *Nature* 417, 835, 2002.
194. Z. F. Ren, Z. P. Huang, J. W. Xu, J. H. Wang, P. Bush, M. P. Siegal, and P. N. Provencio, *Science* 282, 1105 (1998).
195. Y. C. Kong, D. P. Yu, B. Zhang, W. Fang, and S. Q. Feng, *Appl. Phys. Lett.* 78, 407 (2001).
196. Li Mengke, W. Chengwei, and L. Hulin, *Chin. Sci. Bull.* 46, 1793 (2001).
197. D. Routkevitch, A. A. Tager, J. Haruyama, D. Almawlawi, M. Moskovits, and J. M. Xu, *IEEE Trans. Electron Devices* 43, 1646 (1996).
198. Z. Zhang, J. Y. Ying, and M. S. Dresselhaus, *J. Mater. Res.* 13, 1745 (1998).
199. K. Dickmann, J. Jersch, and F. Demming, *Surf. Interface Anal.* 25, 500 (1997).
200. X. Zhang, C. P. Grigoropoulos, D. J. Krajnovich, and A. C. Tam, *IEEE Trans. Compon. Packag. Manuf. Technol. C* 19, 201 (1996).
201. H. J. Mamin and D. Rugar, *Appl. Phys. Lett.* 61, 1003 (1992).
202. K. Hamad-Schifferli, J. J. Schwartz, A. T. Santos. S. Zhang, and J. M. Jacobson, *Nature* 415, 152 (2002).
203. D. G. Cahill, H. E. Fischer, T. Klitsner, E. T. Swartz, and R. O. Pohl, *J. Vac. Sci. Technol. A* 7, 1259 (1989).
204. I. Hatta, *Int. J. Thermophys.* 11, 293 (1990).
205. D. G. Cahill, K. Goodson, and A. Majumdar, *Trans. ASME, J. Heat Transf.* 124, 2 (2002).
206. D. G. Cahil, *Rev. Sci. Instrum.* 61, 802 (1990).
207. T. Borca-Tasciuc, J. L. Liu, T. Zeng, W. L. Liu, D. W. Song, C. D. Moore, G. Chen, K. L. Wang, M. S. Goorsky, T. Radetic, and R. Gronsky, in "Proceedings of the ASME Heat Transfer Division 1999" (L. C. Witte, Ed.), Vol. 364, p. 117. American Society of Mechanical Engineers, New York, 1999.
208. A. R. Kumar, D.-A. Achimov, T. Zeng, and G. Chen, in "Proceedings of the ASME Heat Transfer Division 2000" (J. H. Kim, Ed.), Vol. 366-2, p. 393. American Society of Mechanical Engineers, New York, 2000.
209. T. Borca-Tasciuc, R. Kumar, and G. Chen, *Rev. Sci. Instrum.* 72, 2139 (2001).
210. D. G. Cahill, H. E. Fisher, T. Klitsner, E. T. Swartz, and R. O. Pohl, *J. Vac. Sci. Technol.* 7, 1259 (1989).
211. K. E. Goodson, M. I. Flik, L. T. Su, and D. A. Antoniadis, *Trans. ASME, J. Heat. Transf.* 116, 317 (1994).
212. F. Volklein and J. E. Kessler, *Exp. Tech. Phys.* 33, 343 (1985).
213. P. Nath, K. L. Chopra, *Thin Solid Films* 18, 29 (1973).
214. A. Jacquot, W. L. Liu, G. Chen, J.-P. Fleurial, A. Dauscher, and B. Lenoir, in "Proceedings ICT2002, the 21st Intlernational Conference on Thermoelectrics," p. 353, IEEE, New York, 2002.
215. Z. Zhang and C. P. Grigoropoulos, *Rev. Sci. Instrum.* 66, 115 (1995).
216. T. Borca-Tasciuc, W. L. Liu, J. L. Liu, K. L. Wang, and G. Chen, in "Proceedings of the 2001 National Heat Transfer Conference" (CD-ROM).
217. F. Volklein, *Thin Solids Films* 188, 27 (1990).
218. K. Kurabayashi, M. Asheghi, M. Touzelbaev, and K. E. Goodson, *J. Microelectromech. Syst.* 8, 180, (1999).
219. A. Rosencwaig, J. Opsal, W. L. Smith, and D. L. Weillenborg, *Appl, Phys. Lett.* 46, 1013 (1985).
220. N. M. Amer, M. A. Olmstead, D. Fournier, and A. C. Boccara, *J. Phys. Colloq.* 44, 317 (1983).
221. A. Rosencwaig and A. Gersho, *J. Appl. Phys.* 47, 64 (1976).
222. S. O. Kanstad and P. E. Nordal, *Can. J. Phys.* 64, 1155 (1986).
223. A. C. Boccara, D. Fournier, and J. Badoz, *Appl. Phys. Lett.* 36, 130 (1980).
224. L. Pottier, *Appl. Phys. Lett.* 64, 1618 (1994).
225. J. Hartmann, P. Voigt, and M. Reichling, *J. Appl. Phys.* 81, 2966 (1997).
226. C. A. Paddock and G. Eesley, *J. Appl. Phys.* 60, 285 (1986).
227. K. E. Goodson, O. W. Kading, M. Rosler, and R. Zachai, *J. Appl. Phys.* 77, 1385 (1995).
228. L. J. Inglehart, A. Broniatowski, D. Fournier, A. C. Boccara, and F. Lepoutre, *Appl. Phys. Lett.* 56, 1749 (1990).
229. R. M. Costescu, M. A. Wall, and D. G. Cahill, *Phys. Rev. B* 67, 054302 (2003).
230. O. M. Wilson, X. Hu, D. G. Cahill, and P. V. Braun, *Phys. Rev. B* 66, 224301 (2002).
231. G. Chen, C. L. Tien, X. Wu, and J. S. Smith, *Trans. ASME, J. Heat Transf.* 116, 325 (1994).
232. T. Borca-Tasciuc and G. Chen, *Intl. J. Thermophys.* 19, 557, 1998.
233. D.-A. Borca-Tasciuc, G. Chen, Y. M. Lin, O. Rabin, M. S. Dresselhaus, A. Borshchevsky, J. P. Fleurial, and M. A. Ryan, in "Nanophase and Nanocomposite Materials IV" (S. Komarneni, R. A. Vaia, G. Q. Lu, J.-I. Matsushita, and J. C. Parker, Eds.), Vol. 703, p. V2.7.1-6. Material Research Society, Warrendale, PA, 2001.
234. S. W. Indermuehle and R. B. Peterson, *Trans. ASME, J. Heat Transf.* 121, 528 (1999).
235. L. Shi, Mesoscopic Thermophysical Measurements of Microstructures and Carbon Nanotubes, Ph.D Thesis, University of California at Berkeley, 2001.
236. M. Nonnenmacher and H. K. Wickramasinghe, *Appl. Phys. Lett.* 61, 168 (1992).
237. K. Luo, Z. Shi, J. Varesi, and A. Majumdar, *J. Vac. Sci. Technol. B* 15, 349 (1997).
238. R. J. Pylkki, P. J. Moyer, and P. E. West, *Jpn. J. Appl. Phys. 1* 33, 3785 (1994).
239. M. Maywald, R. J. Pylkki, and L. Balk, *Scanning Microsc.* 8, 181 (1994).
240. A. Majumdar, *Annu. Rev. Mater. Sci.* 29, 505 (1999).
241. E. Betzig and J. K. Trautman, *Science* 257, 189 (1992).
242. B. D. Boudreau, J. Raja, R. J. Hocken, S. R. Patterson, and J. Patten, *Rev. Sci. Instrum.* 68, 3096 (1997).
243. K. E. Goodson and M. Asheghi, *Microscale Thermophys. Eng.* 1, 225 (1997).
244. D. A. Fletcher, K. B. Crozier, C. F. Quate, G. S. Kino, K. E. Goodson, D. Simanovskii, and D. V. Palanker, *Appl. Phys. Lett.* 77, 2109 (2000).
245. A. M. Mansanares, D. Fournier, and A. C. Boccara, *Electron. Lett.* 29, 2045 (1993).
246. J. Christofferson, D. Vashaee, A. Shakouri, and P. Melese, *Proc. SPIE* 4275, 119 (2001).
247. S. Ju, Q. W. Kading, Y. K. Leung, S. S. Wong, and K. E. Goodson, *IEEE Electron Device Lett.* 18, 169 (1997).
248. H. Brugger, P. W. Epperlein, S. Beeck, and G. Abstreiter, in "Proceedings of the 16th International Symposium Gallium Arsenide and Related Compounds Conference" (T. Ikoma and H. Watanabe, Eds.), Vol. 24, p. 771. IOP, Bristol, UK, 1990.
249. R. Puchert, A. Barwolff, U. Menzel, A. Lau, M. Voss, and T. Elsaesser, *J. Appl. Phys.* 80, 5559 (1996).
250. U. Menzel, R. Puchert, A. Barwolff, and A. Lau, *Microelectron. Reliab.* 38, 821 (1998).
251. V. Spagnolo, M. Troccoli, G. Scamarcio, C. Becker, G. Glastre, and C. Sirtori, *Opt. Mater.* 17, 263 (2001).
252. Y. Imry and R. Landauer, *Rev. Mod. Phys.* 71, S306 (1999).
253. M. F. Modest, "Radiative Heat Transfer." McGraw–Hill, New York, 1993.
254. D. K. Ferry, "Semiconductors." MacMillan Co., New York, 1991.
255. J. Callaway, *Phys. Rev.* 113, 1046 (1959).
256. C. Cercignani, "Rarefied Gas Dynamics: From Basic Concepts to Actual Calculations." Cambridge, New York, 2000.
257. H. B. Casimir, *Physica* 5, 495 (1938).
258. F. Volklein and E. Kessler, *Thin Solid Films* 142, 169 (1986).
259. U. Dilliner and F. Volklein, *Thin Solid Films* 187, 263 (1990).

260. G. Chen and C. L. Tien, *AIAA J. Thermophys. Heat Transf.* 7, 311 (1993).
261. S. Kumar, G. C. Vradis, *Trans. ASME, J. Heat Transf.* 116, 28 (1994).
262. M. I. Flik and C. L. Tien, *Trans. ASME, J. Heat Transf.* 112, 872 (1990).
263. A. Majumdar, *Trans. ASME, J. Heat Transf.* 115, 7 (1993).
264. C. P. Jen and C. C. Chieng, *J. Thermophys. Heat Transf.* 12, 146 (1998).
265. T. Zeng and G. Chen, *Trans. ASME, J. Heat Transf.* 123, 340 (2001).
266. X. Y. Yu, G. Chen, A. Verma, and J. S. Smith, *Appl. Phys. Lett.* 67, 3554 (1995).
267. C. K. Chan and C. L. Tien, *Trans. ASME, J. Heat Transf.* 95, 302 (1973).
268. D. W. Song, W.-N. Shen, T. Zeng, W. L. Liu, G. Chen, B. Dunn, C. D. Moore, M. S. Goorsky, R. Radetic, and R. Gronsky, in "Proceedings of the ASME Heat Transfer Division 1999" (L. C. Witte, Ed.). American Society of Mechanical Engineers, New York, 1999.
269. G. Gesele, J. Linsmeier, V. Drach, J. Fricke, and R. Arens-Fischer, *J. Phys. D* 30, 2911 (1997).
270. V. Lysenko, V., Gliba, V. Strikha, A. Dittmar, G. Delhomme, Ph. Roussell, D. Barbier, N. Jaffrezic, and C. Martelet, *Appl. Surf. Sci.* 123, 458 (1998).
271. J. D. Chung and M. Kaviany, *Int. J. Heat Mass Tranf.* 43, 521 (2000).
272. J. O. Sofo and G. D. Mahan, in "18th International Conference on Thermoelectrics: Proceedings, ICT'99," Vol. 736, p. 626. IEEE, Piscataway, NJ, 1999.
273. Ce-Wen Nan, R. Birringer, David R. Clarke, and H. Gleiter, *J. Appl. Phys.* 81, 6692 (1997).
274. O. Tornblad, P. G. Sverdrup, D. Yergeau, Z. Yu, K. E. Goodson, and D. W. Dutton, in "2000 International Conference on Simulation Semiconductor Processes and Devices," Vol. 282, p. 58. IEEE, Piscataway, NJ, 2002.
275. J. R. Howell, *J. Heat Transf.* 120, 547 (1998).
276. C. Moglestue, *Comput. Methods Appl. Mech. Eng.* 30, 173 (1982).
277. C. Jacoboni and L. Reggiani, *Rev. Mod. Phys.* 55, 645 (1983).
278. M. V. Fischetti and S. E. Laux, *Phys. Rev. B* 38, 9721 (1988).
279. M. V. Fischetti and S. E. Laux, *Phys. Rev. B* 48, 2244 (1993).
280. P. Lugli, P. Bordone, L. Reggiani, M. Rieger, P. Kocevar, and S. M. Goodnick, *Phys. Rev. B* 39, 7852 (1989).
281. R. B. Peterson, *Trans. ASME, J. Heat Transf.* 16, 815 (1994).
282. T. Klitsner, J. E. VanCleve, H. E. Fischer, and R. O. Pohl, *Phys. Rev. B* 38, 7576 (1988).
283. S. Mazumder and A. Majumdar, *Trans. ASME, J. Heat Transf.* 123, 749 (2001).
284. R. L. Liboff, "Kinetic Theory." Prentice Hall, Englewood Cliffs, NJ, 1998.
285. G. D. Mahan, "Many-Particle Physics." Kluwer Academic/Plenum P. New York, 2000.
286. Z. B. Zhang, M. S. Dresselhaus, and J. Y. Ying, *Mater. Res. Soc. Symp. Proc.* 524, 351 (1999).
287. Y. M. Lin, S. B. Cronin, J. Y. Ying, J. Heremans, and M. S. Dresselhaus, in *Mater. Res. Soc. Symp. Proc.* 635, C4.30.1 (2001).
288. J. L. Liu, J. Wan, Z. M. Jiang, A. Khitun, K. L. Wang, and D. P. Yu, *J. Appl. Phys.* 92, 6804 (2002).
289. B. Pendry, "Low Energy Electron Diffraction." Academic Press, New York, 1974.
290. M. V. Simkin and G. D. Mahan, *Phys. Rev. Lett.* 84, 927 (2000).
291. B. Yang and G. Chen, in "Proceedings ICT2002: 21st International Conference on Thermoelectrics," p. 306, IEEE, New York, 2002.
292. D. Frenkel and B. Smit, "Understanding Molecular Simulation, From Algorithm to Applications." Academic Press, London, 1996.
293. D. C. Rapaport, "The Art of Molecular Dynamics." Cambridge Univ. Press, Cambridge, UK, 1995.
294. J. P. Hansen and I. A. MacDonald, "Theory of Simple Liquids." Academic Press, New York, 1990.
295. Q. M. Yu, M. O. Thompson, and P. Clancy, *Phys. Rev. B* 53, 8386 (1996).
296. S. Kotake and S. Wakuri, *JSME Int. J. B* 37, 103 (1994).
297. S. Volz, J. B. Saulnier, M. Lallemand, B. Perrin, P. Depondt, and M. Mareschal, *Phys. Rev. B* 54, 340 (1996).
298. R. D. Mountain and R. A. MacDonald, *Phys. Rev. B* 28, 3022 (1983).
299. A. Tenenbaum, G. Ciccotti, and R. Gallico, *Phys. Rev. A* 25, 2778 (1982).
300. R. H. H. Poetzsch and H. Boettger, *Phys. Rev. B* 50, 15757 (1994).
301. S. Volz and G. Chen, *Phys. Rev. B* 61, 2651 (2000).
302. A. J. C. Ladd, B. Moran, and W. G. Hoover, *Phys. Rev. B* 34, 5058 (1986).
303. J. Che, T. Cagin, W. Deng, and W. A. Goddard III, *J. Chem. Phys.* 113, 6888 (2000).
304. J. R. Lukes, D. Y. Li, X. G. Liang, and C. L. Tien, *Trans. ASME, J. Heat Transf.* 122, 536 (2000).
305. L. J. Porter, J Li, and S. Yip, *J. Nucl. Mater.* 246, 53 (1997).
306. J. Li, L. J. Porter, and S. Yip, *J. Nucl. Mater.* 255, 139 (1998).
307. F. H. Stillinger and T. A. Webber, *Phys. Rev. B* 31, 5262 (1985).
308. M. Moseler and U. Landman, *Science* 289, 1165 (2000).
309. A. Nakano, M. E. Bachlechner, P. Branicio, T. J. Campbell, I. Ebbsjo, R. K. Kalia, A. Madhukar, S. Ogata, A. Omeltchenko, J. P. Rino, F. Shimojo, P. Walsh, and P. Vashishta *IEEE Trans. Electron Devices* 47, 1804 (2000).
310. T. Ikeshoji and B. Hafskjold, *Mol. Phys.* 81, 251 (1994).
311. F. Muller-Plathe, *J. Chem. Phys.* 106, 6082 (1997).
312. J. R. Lukes, D. Y. Li, X. G. Liang, and C. L. Tien, *Trans. ASME, J. Heat Transf.* 122, 536 (2000).
313. R. Kubo, *J. Phys. Soc. Jpn.* 12, 570 (1957).
314. R. Kubo and M. Yokota, *J. Phys. Soc. Jpn.* 12, 1203 (1957).
315. R. Zwanzig, *Anna. Rev. Phys. Chem.* 16, 67 (1964).
316. S. Volz, J. B. Saulnier, G. Chen, and P. Beauchamp, *High Temp.–High Press.* 32, 709 (2000).
317. X. G. Liang and B. Shi, *Mater. Sci. Eng. A* 292, 198 (2000).
318. P. Chantrenne and M. Raynaud, in "Proceedings of the 12th International Heat Transfer Conference" (J. Taine, Ed.), Vol. 1, p. 549. Elsevier, New York, 2002.
319. A. R. Abramson, C.-L. Tien, and A. Majumdar, *Trans. ASME, J. Heat Transf.* 124, 963 (2002).
320. Chen, Voltz, and Freund, private communication.
321. R. J. Stoner and H. J. Maris, *Phys. Rev. B* 48, 16373 (1993).
322. M. E. Lumpkin and W. M. Saslow, *Phys. Rev. B* 20, 1035 (1979).
323. Y. Ge and S. Chen, *Solid State Commun.* 77, 313 (1991).
324. A. Maiti, G. D. Mahan, and S. T. Pantelides, *Solid State Commun.* 102, 517 (1997).
325. S. M. Rytov, Theory of Electric Fluctuations and Thermal Radiation, AFCRC-TR-59-162, Air Force Cambridge Research Center, Bedford, MA, 1959.
326. G. Chen, *Microscale Thermophys. Eng.* 1, 215 (1997).
327. L. M. Phinney and C. L. Tien, *Trans. ASME, J. Heat Transf.* 3, 751 (1998).
328. K. Blotekjaer, *IEEE Trans. Electron Devices* 17, 38 (1970).
329. K. A. Fushinobu, A. Majumdar, and K. Hijikata, *Trans. ASME, J. Heat Transf.* 117, 25 (1995).
330. A. Majumdar, K. A. Fushinobu, and K. Hijikata, *J. Appl. Phys.* 77, 6686 (1995).
331. J. Lai and A. Majumdar, *Appl. Phys.* 79, 7353 (1996).
332. T. Q. Qiu and C. L. Tien, *Int. J. Heat Mass Transf.* 37, 2799 (1994).
333. M. Lundstrom, "Fundamentals of Carrier Transport," 2nd ed. Cambridge Univ. Press, Cambridge, UK, 2000.

334. H. J. Goldsmid, "Thermoelectric Rrefrigeration." Plenum Press, New York, 1964.
335. L. P. Bulat and V. G. Yatsyuk, *Sov. Phys.-Semicond.* 18, 383 (1984).
336. L. I. Anatychuk, L. P. Bulat, D. D. Nikirsa, and V. G. Yatsyuk, *Sov. Phys.-Semicond.* 21, 206 (1987).
337. G. Chen, *Int. J. Therm. Sci.* 40, 693 (2001).
338. R. G. Yang and G. Chen, in "Proceedings of the International Mechanical Engineering Conference and Exhibitions," 2001 (CD-ROM).
339. S. V. J. Narumanchi, J. Murthy, and C. H. Amon, in "Proceedings of the International Mechanical Engineering Conference and Exhibitions," 2001 (CD-ROM).
340. G. Chen, *Appl. Phys. Lett.* 82, 991 (2003).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanoscale Magnetic Random Access Memory Elements

S. J. Pearton

University of Florida, Gainesville, Florida, USA

J. R. Childress

IBM Almaden Research Center, San Jose, California, USA

CONTENTS

1. INTRODUCTION

Ferromagnetic thin films and multilayers have been researched intensively in recent years for applications in various magnetic recording and nonvolatile memory devices [1]. Interest in these materials has increased dramatically since the discovery of giant magnetoresistance (GMR) in multilayers comprised of alternating ultrathin (10 to 50 Å) ferromagnetic/noble metal layers. Briefly, the GMR effect can be understood in terms of spin-dependent scattering of conduction electrons at ferromagnetic/nonmagnetic interfaces. Conduction electrons with a spin direction parallel to a material's magnetic orientation move freely, providing low resistance. Conversely, conduction electrons with spin direction opposite to the material's magnetic orientation are hampered by more frequent collisions with atoms in the material, producing higher resistance.

In computer and data processing systems the main form of data storage and retrieval is based on magnetic recording systems, either magnetic disks or tape drives [2–16]. Information is written and stored as magnetization patterns on a recording media and can be transferred back and forth using a magnetic sense head. In addition magnetic random access memories (MRAM) are used for storage and processing of very high bit densities. These devices offer advantages over semiconductor memories as they are radiation-hard and infinitely rewritable, with high storage densities, and fast access time. Finally there are magnetic sensors, transducers, and actuators which are used in automobiles, aircraft, hydraulic equipment, and defense applications (mine detection, perimeter defense) [1, 9]. In all of these structures there is the need to pattern the magnetic layers [17] which are generally thin films of materials such as NiFe, NiFeCo, which are often incorporated into multilayers comprising magnetic and nonmagnetic materials.

The fabrication of read/write heads and other magnetic storage elements requires methods for producing submicrometer features in multilayer structures involving NiFe and NiFeCo [1, 13–17]. Due to the relative involatility of the etch products of these materials in conventional plasma reactors, virtually all of the patterning is performed using ion beam etching or lift-off processes [17–20]. Both of these methods have drawbacks for producing very small features—ion beam etching has poor selectivity with respect to mask materials and can produce redeposition of the sputtered materials onto the feature sidewalls that degrades the magnetic performance of the devices [18, 19] while lift-off processes typically suffer from poor yield relative to direct etch removal methods.

In a typical reactive plasma process, exposure of NiFe and NiFeCo to chlorine or other associated feed gases produces a reaction or selvedge layer consisting of involatile etch products (i.e., $NiCl_X$, $FeCl_X$, $CoCl_X$ species) [18–22]. Since these do not leave the surface and their atomic volumes are typically larger than those of the elements they are formed from, there is actually net deposition, rather than etching. There are basically two methods to enhance desorption of the etch products. The first is to heat the sample during plasma exposure to increase etch product volatility, but typically temperatures above 300 °C are required [18]

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (461–475)

and this exceeds the thermal "comfort zone" of most magnetic multilayer structures. The second is to employ a high ion flux in conjunction with the reactive chlorine neutral flux incident on the sample. The ion energy, however, must be kept low under these conditions to avoid mask degradation and loss of etch selectivity. The etch rates are a strong function of ion flux, ion energy, and plasma gas composition, all of which may be interpreted in terms of balancing formation of chloride etch products with concurrent ion-assisted desorption of these products. Thus, the role of ion bombardment is critical in this mechanism. Alternative plasma chemistries including bromine or iodine containing plasmas were also examined under high density plasma conditions. The interhalogens ICl and IBr have been found to dissociate readily in high density plasma sources, producing high concentrations of reactive species.

However, one concern with this process is the effect of residual chlorine or chlorinated etch residues remaining on the sidewalls of etched features. We report on the effectiveness of several different *in-situ* or *ex-situ* cleaning processes after Cl_2/Ar high density plasma etching of GMR elements. We find that inductively coupled plasma (ICP) etching does not degrade the magnetic performance of single or multilayer structures, and subsequent water rinsing or *in-situ* H_2 or SF_6 plasma cleaning is efficient in preventing corrosion.

Another method for preventing corrosion is applying noncorrosive gas chemistries such as CO/NH_3 or CO_2/NH_3. Nakatani [23] reported use of a noncorrosive CO/NH_3 plasma chemistry for NiFe, in which the etch products are expected to be carbonyls [e.g., $Fe(CO)_5$]. The resultant etch rates were $\leq$300 Å/min^{-1} for $Ni_{0.8}Fe_{0.2}$, a factor of about three higher than purely physical Ar^+ sputtering under the same conditions. The reactor employed was noncommercial and might be classified as a magnetron-type, medium ion density system. We report on a parametric investigation of CO/NH_3 etching of NiFe, NiFeCo, and related thin film materials using an ICP reactor. We find maximum etch rates for NiFe and NiFeCo in the range 350–400 Å/min^{-1}, with the rates being a strong function of $CO:NH_3$ ratio, source, and radio frequency chuck power, and pressure. We also compared use of CO_2 to replace CO, though this does not appear to offer any advantage in etch rates.

In this chapter, a comparison is given of high density plasma reactors such as electron cyclotron resonance, and ICP based on either corrosive gas chemistries (Cl_2, IBr, ICl) or noncorrosive gas chemistries (CO/NH_3, CO_2/NH_3) will be described for the etching of NiFe, NiFeCo, and other elements in magnetic multilayers. The influence of postetch cleaning procedures on magnetic properties is also examined. Finally, the selectivity for etching the magnetic materials over common mask materials (photoresist, SiO_2) was measured over a broad range of conditions.

2. MAGNETIC DEVICES

2.1. Theory of GMR

In any material, some conduction electrons will be oriented with "up" spins, and some will be oriented with "down" spins. The material can be viewed as having two independent conduction channels, one for each electron spin orientation. The measured current will be the sum of these channels:

$$j = \sum_s j_s \qquad (s = \uparrow, \downarrow) \tag{1}$$

In a nonmagnetic material, these two currents are equal. In a magnetic material, such as Fe or Co, the two currents are different: interaction of a conduction channel with its surroundings will be spin-dependent. First, the conduction electron energy bands will be different for majority spin electrons (spin parallel to the magnetization). Second, minority and majority spin electrons will have different scattering potentials when they encounter impurities and interfaces. Thus, in a bulk ferromagnetic material, the resistivity of the majority spin channel ($\rho\uparrow$) is typically different from the minority spin channel ($\rho\downarrow$) [24, 25].

In a multilayer structure where the layers are antiparallel, however, the up conduction channel will be scattered as majority spin electrons in one layer and as minority spin electrons in the next layer, and vice versa. For this reason, the net resistivities of two channels will be the same and approximately equal to the average resistivity of the two channels in the bulk:

$$\begin{aligned} \rho_\uparrow &= a + b \\ \rho_\downarrow &= a - b \\ \bar{\rho} &\cong \frac{1}{2}(\rho_\uparrow + \rho_\downarrow) = a \end{aligned} \tag{2}$$

Conductivities (σ) of the two channels (σ_1, σ_2) add in parallel and the resistivity is the reciprocal of the conductivity. So

$$\rho_{\text{tot}} = \frac{1}{\sigma_1 + \sigma_2} = \frac{\rho_1 \rho_2}{\rho_1 + \rho_2} \tag{3}$$

In the absence of an applied field, $\rho_1 = \rho_2 = \bar{\rho}$, and the total resistivity will be

$$\rho_0 = \frac{1}{2}\bar{\rho} = \frac{1}{2}a \tag{4}$$

At the saturation field, on the other hand, $\rho_1 = \rho\uparrow$ and $\rho_2 = \rho\downarrow$. The layers will be oriented in parallel:

$$\rho_{\text{sat}} = \frac{\rho_\uparrow \rho_\downarrow}{\rho_\uparrow + \rho_\downarrow} = \frac{a^2 - b^2}{2a} \tag{5}$$

Clearly,

$$\rho_{\text{sat}} < \rho_0 \tag{6}$$

In other words, an applied magnetic field (for instance, a field stored in magnetic media) "short circuits" the magnetic layers, reducing the total resistivity. This short circuit is observed as the GMR effect.

2.2. Basic Mechanism of GMR

Giant magnetoresistance is largely caused by a quantum effect called spin-dependent scattering, which results from the fact that the spin of an electron can point either up or down. Electrical resistance is caused by the scattering of electrons, but magnetic materials scatter spin-up and spin-down electrons differently (Fig. 1) [26].

Given a difference in resistivity between spin-up and spin-down electrons, the low resistance of the multilayer is generated when conduction electrons with a spin direction parallel to a material's magnetic orientation move freely. Conversely the higher resistance can be produced when conduction electrons with spin direction opposite to the material's magnetic orientation are hampered by more frequent collisions with atoms in the material.

A structure consists of two magnetic layers separated by a nonmagnetic spacer layer. If the layers are magnetized in the same direction, both will scatter electrons in the same way and the structure will have almost the same resistance as the bulk material. However, if the layers are magnetized in opposite directions, one layer will mainly scatter spin-up electrons, while the other will scatter spin-down electrons. The overall resistance will increase.

The key point is that only a weak magnetic field is needed to change the orientation of the magnetic fields in the layers. By using a strong antiferromagnet to keep the orientation in one of the layers constant, changes in resistance can be linked to the magnetic state of the other layer.

There are two significant advantages of GMR over competing devices. First, the large resistance change yields a strong signal and, second, the technology is compatible with integrated circuit technology, so that GMR devices can be included as part of a chip package to make smaller, faster, less expensive sensors and memory chips.

2.3. GMR Read Heads

In the past couple of years, production of read heads for hard-disk drives has been largely converted from thin-film inductive heads to magnetoresistance heads using thin-film 80:20 nickel/iron magnetoresistors. This conversion was needed in order to get larger signals from smaller storage bits and keep pace with improvements in hard drives.

A read head detects a transition in magnetization stored in the disk as the head passes over it. The magnetic field produced by the transition is either up (head-to-head magnetization transition) or down (tail-to-tail magnetization transition).

The GMR read head uses a feature called a spin valve, which is etched into a material whose edge is oriented along the direction of the disk surface. Spin valve material is normally composed of four thin films: a sensing layer, a conducting spacer, a pinned layer, and an exchange layer. The first three films are very thin, allowing conduction electrons to frequently move back and forth between the sensing and pinned layers via the conducting spacer (Fig. 2) [26]. The magnetic orientation of the pinned layer is fixed and held in place by the adjacent exchange layer, while the magnetic orientation of the sensing layer changes in response to the magnetic field from the disk. A change in the magnetic orientation of the sensing layer will cause a change in the resistance of the combined sensing and pinned layers.

In the GMR read head, the magnetization of the pinned spin-valve layer is directed vertical to the disk surface, and the soft layer lies parallel to the surface in the absence of a field from the disk. Up and down magnetic fields, which create smaller or larger resistance in the spin valve, are generated by stored magnetic data on a rotating disk.

2.4. Magnetic Random Access Memory

Nonvolatility, the ability to store data when electricity is off, is a much desired memory property in applications where data retention is critical. Semiconductor nonvolatile random access memory technologies, such as electrically erasable programmable read-only memory and flash memory, suffer from slow write times and wear out after data are stored more than 1 million times. Memory that uses magnetic materials can have fast write times and can store data indefinitely.

Working memory systems using these MRAM chips have been used for space and missile applications in which resistance to radiation damage is critical. Nonvolatility and

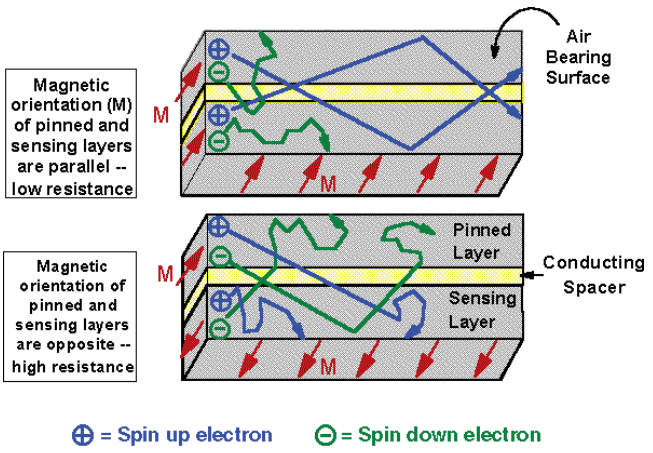


Figure 1. The giant magnetoresistance effect is due to the large difference in electrical resistance between two magnetic states of a metallic multilayer film.

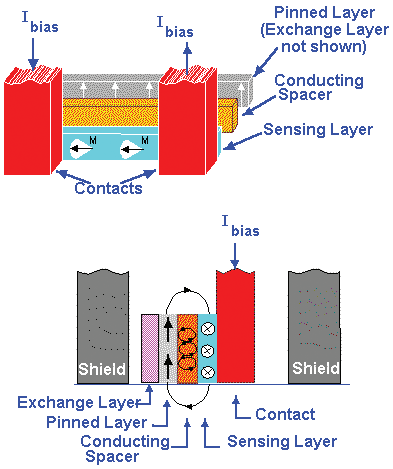


Figure 2. Schematic and operation of GMR head.

radiation hardness are important for space missions, and magnet storage is intrinsically radiation-hard. Recently, GMR materials have been used to make MRAMs with faster read access times.

With proper design of the areal geometry and the right thickness of magnetic films, submicrometer-size MRAM cells (single memory bits) have been made to operate like a spin valve, and they have been called "pseudo spin valve" devices. They promise very high density, high speed, and nonvolatile memory applications.

A newer high-magnetostrictive innovation, called a spin-dependent tunneling (SDT) device or magnetic tunnel junction, uses tunneling current through a thin dielectric between two ferromagnetic films. SDT devices work in a way similar to GMR devices. When two magnetic layers are magnetized in the same direction, the tunneling current is generally higher than when the two magnetic devices are magnetized in opposite directions.

3. PATTERNING PROCESSES

3.1. Transition from Wet Etching to Dry Etching

Etch processes may be classified by their rate, selectivity, uniformity, directionality (isotropy or anisotropy), surface quality, and reproducibility. All etching processes involve three basic events:

1. Movement of the etching species to the surface to be etched.
2. Chemical reaction to form a compound that is soluble in the surrounding medium.
3. Movement of the by-products away from the etched region, allowing fresh etchant to reach the surface.

Both (1) and (3) usually are referred to as diffusion, although convection may be present. The slowest of these processes primarily determines the etch rate, which may be diffusion or chemical-reaction limited. There are two different etching methods by using two quite different media: liquid chemicals (wet etching) and reactive gas plasmas (dry etching). Wet etching is performed by immersing the wafers in an appropriate solution or by spraying the wafer with the etchant solution [27].Wet-chemical etching is superior to dry etching in terms of effectiveness, simplicity, low cost, low damage to the wafer, high selectivity, and high throughput. However, the main limitations of wet etching include its isotropic nature which results in roughly equal removal of material in all directions, making it incapable of patterning submicrometer features, and the need for disposal of large amounts of corrosive and toxic materials. As the requirements developed for increasing circuit density and narrower linewidth in the manufacture of very large-scale integrated circuit/ultra-large-scale integrated circuit devices, it became necessary to have new etching methods to replace the wet etching. Dry etching methods became favorable etching processes for integrated circuit manufacture. Plasma-driven chemical reactions and/or energetic ion beams are used to remove materials in dry etching system. The most significant advantage of dry over wet etching is that it provides higher resolution potential by overcoming the problem of isotropy. Other benefits are the reduced chemical hazard and waste treatment problems, and the ease of process automation and tool clustering.

3.2. Ultimate Goals of a Dry Etch Process

The success of a etch process must be measured by the following nine parameters. The greatest challenge is that each parameter can usually only be optimized at the expense of at least one of the others.

3.2.1. Critical Dimension Uniformity

Uniformity across the wafer—including densely populated areas and large open spaces and within high aspect ratio features—is critical to maintain consistent device performance. Aspect ratio dependent etching, also known as "microloading," is a common nonuniform problem.

3.2.2. Selectivity

Defined as the ratio of the etch rate of one material versus that of another, the selectivity of the material being etched to the overlying masking layer (typically photoresist) is usually of the most concern, since this impacts critical dimension and profile control and the thickness of resist required (thinner photoresist is required to adequately resolve smaller feature sizes, so selectivity must increase as geometry shrinks). Also of concern is the selectivity to the underlying material upon which the etch stops. Different selectivity specifications may be given for edges and flat areas, since edges tend to etch faster.

3.2.3. Etch Rate

A high etch rate is needed to keep the throughput of the system or process module high (usually measured in Å/min). There is usually a trade-off between etch rate and other parameters, such as selectivity and damage.

3.2.4. Etch Profile Control

It is usually desirable to have an anisotropic profile, that is, one where the etched feature edges are close to vertical, to maximize packing density on the chip. But it is also desirable to have a flare out at the top of the feature to enable good step coverage in subsequent deposition steps.

3.2.5. Low Damage

Damage is an obvious concern. The high energy of the plasma can create currents on the wafer surface that cause electrical damage and energetic ions can cause mechanical damage to the films' crystalline structure.

3.2.6. Residue

Residue which coats the interior of the etch chamber is a difficult problem to avoid. In addition to requiring more frequent cleaning, residue is also a source of contamination. The most significant factors in controlling residue are temperature, bottom rf power, backside cooling, and process pressure.

3.2.7. Corrosion

Corrosion is mainly a problem in metal etch. Upon exposure to water vapor (i.e., air), chlorine will immediately attack metals. Integrated postetch treatments help eliminate this problem.

3.2.8. Particle Control

Particle control is another critical measurement of etch system performance. Today, fewer than 0.05 particles/cm^2 that are >0.35 μm in size are required.

3.2.9. Sidewall Passivation

Sidewall passivation is important both during and after the etch. Carbon from the photoresist mask typically combines with etching gases and etch by-products to form a polymerlike material on the sidewall of the feature. This is usually a requirement in creating anisotropic profiles. The biggest challenge is that after the etch, this polymer must be removed.

3.3. Basic Mechanism of the Etch Process

The optimization of etch of these parameters—uniformity, selectivity, etch rate, profile control, damage, and residue control—requires an understanding and fine-tuning of the two very different mechanisms through which etching occurs.

As shown in Figures 3 and 4, one mechanism is purely chemical. Reactive species generated in plasma react with the wafer surface and create volatile etch products that are swept away. By careful selection of the gases that are flowed into the plasma (typically chlorine and/or fluorine containing gases), it is possible to achieve very high selectivity through this process. However, since films tend to etch in all directions at once, the result is an isotropic etch. The other mechanism is purely physical. Energetic ions crossing the sheath transfer large amounts of energy and momentum to the substrate. The force of these ions can be strong enough to physically remove material. At low pressure where the mean free path is long, the ejected sputtered material can cross the reactor vessel and reach opposing walls. The main benefit of this etch mechanism is that it provides some directionality to the etch, making it possible to achieve highly anisotropic profiles. However, it is the least selective mechanism and also suffers from the disadvantages of low etch rate, damage, and trenching [28].

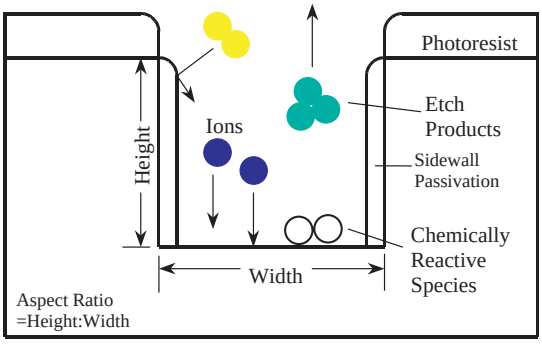


Figure 3. Schematic of plasma etching; etching occurs because of two etch mechanisms: chemical reaction and ion bombardment.

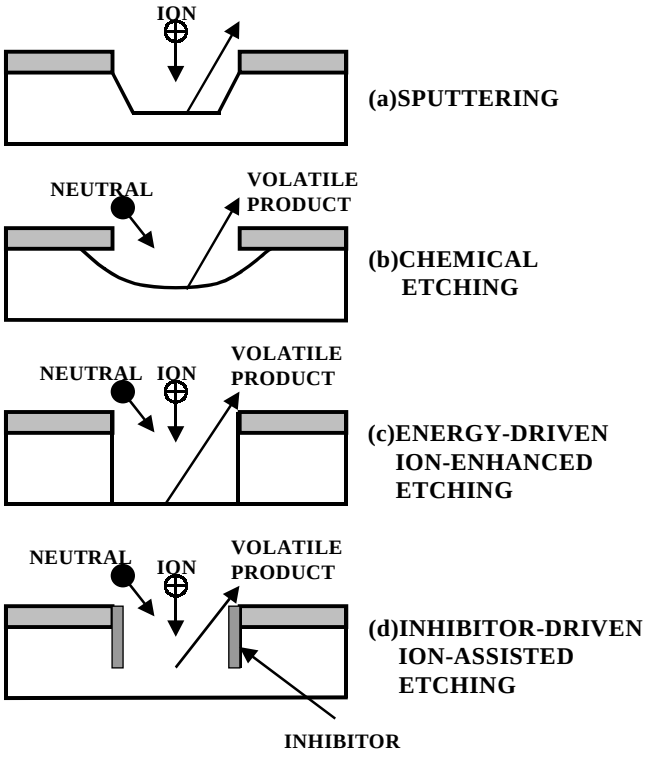


Figure 4. The four basic mechanisms of plasma etching: (a) sputtering, (b) purely chemical etching, (c) energy-driven ion-enhanced etching, and (d) inhibitor-driven ion-assisted etching.

3.3.1. Sputtering

In sputtering, impinging particles (usually positive ions accelerated across the sheath) strike the surface with high kinetic energy. Some of the energy is transferred to surface atoms which then are ejected, leading to a net removal of material. This process is distinguished from other etching mechanisms in that the interaction is mechanical. It is sensitive to the magnitude of bonding forces and structure of a surface, rather than its chemical nature, and quite different materials can sputter at similar rates. In a way this is symptomatic of using ion bombardment with energy far higher than the surface binding energy.

3.3.2. Chemical Etching

Chemical etching comes about when active species from the gas phase encounter a surface and react with it to form a volatile product. Product volatility is necessary for chemical etching since involatile products would coat the surface and protect it from further attack. In this type of mechanism the plasma reactor converts the feed into reactive chemical species, which are usually free radicals. There is usually no directionality and the etching can be specific (high selectivity) since it its governed by the relative chemical affinities between the etchant species and exposed materials. Because of this lack of directionality, chemical etching is commonly called isotropic etching.

3.3.3. Energy-Driven Ion-Enhanced Etching

There is usually little or no etching when the substrate surface is exposed to neutral chemical species alone in the absence of ion bombardment. Impinging ions damage the substrate material by virtue of their impact energy and thereby render the solid substrate more reactive toward incident neutral radicals.

3.3.4. Inhibitor-Driven Ion-Assisted Etching

Inhibitor-protected sidewall ion-enhanced etching differs from energy-driven ion-enhanced etching in that the chemical etching reaction in spontaneous, even without ion bombardment. Neutral etchant species from the plasma spontaneously gasify the substrate, and ions play a role by interaction with another component—a "protective" inhibitor film. The role of ions in the surface-inhibitor mechanism is to clear the inhibitor from horizontal surfaces that are bombarded by the flux of ions impinging in the vertical direction. The protective film is not removed from the vertical walls of masked features because these surfaces only intercept those few ions that are scattered as they cross the sheath. This protective film may originate from involatile etching products or from film-forming precursors that adsorb during the etching process.

3.4. Dry Etching Techniques

3.4.1. Plasma Etching

A wafer is exposed to a reactive gas such as chlorine which is in some cases dissociated in a plasma to create highly reactive atoms. Etching is isotropic or crystallographic, and temperature and reactant flux are used to adjust etch character. During the purely chemical etching process, three steps occur: adsorption of the necessary species on the materials surface, chemical reaction, and desorption of the products. The advantage of the technique is rapid etch rates, but the drawbacks are isotropy, a tendency for strong loading effects, and release of heat.

3.4.2. Ion Beam Etching

Ion beam etching (ion milling) uses a broad-area ion beam composed of a nonreactive gas such as Ar with high ion energy. This technique only relies on physical sputtering. The uniform ion beam bombards a wafer to cause etching. Since ion beam etching is very anisotropic and the wafer is not exposed to plasma, it is well suited for etching of certain patterns.

A major problem with ion beam etching is that the etch rate is very low because of the nature of physical sputtering, and it is also recognized that the etch rate of ion milling is very much dependent upon the incidence angle. It typically peaks at between 30° and 50° and becomes very small when angle is larger than 80°. There are a number of factors that limit the application of ion beam etching. First of all, etch products, usually nonvolatile, can redeposit on the wafers, especially onto the sidewalls of the etch mask that degrades the magnetic performance of the devices. When the film plane is normal to the ion beam, the etch rate of the sidewalls is very small since the incidence angle there is close to 90°. As a result, after the etch mask is removed, undesired fence and trenches are often left on the edges of the etched patterns.

3.4.3. Reactive Ion Etching

Reactive ion etching uses radio frequency (rf) power to maintain a plasma. Applied rf power makes electrons accelerate in the sheath region changing direction upward and downward. The accelerated electrons can lose a large fraction of their kinetic energy through dissociative collisions with molecules and atoms. Because electrons are light and have high energy, they diffuse fastest, leaving an excess of positive charge and a plasma potential that is positive relative to the electrode. Since charged particles are most abundant in the central glow of the plasma, most of the potential drop appears across the sheath. Positive ions are accelerated through the sheath and strike the samples, giving rise to a physical etch component. In addition, ion flux is coupled with ion energy. High plasma densities and ion fluxes are gained at the expense of extremely high applied voltages and damage levels.

3.5. High-Density Plasma Reactor

Increasingly, the limitation of traditional plasma etch technologies—which has resulted in the trend to high density plasma sources—is primarily one of process pressure. In the pressure regime of a few hundred millitorr where it is relatively straightforward to create a plasma, it becomes difficult to get etchant in and reaction by-products out of openings that are smaller than about 0.25 μm. The problem is more severe with higher aspect ratio.

The solution is to go to lower pressures, where the mean free path lengths of gas molecules and ions are longer, which reduces scattering collisions that can cause loss of profile control.

However, this is not as simple as it may seem since it requires a switch to a different type of plasma source—a so-called high density source—that is capable of generating enough ions to achieve acceptable etch rates at reduced pressure. High density sources are designed to more efficiently couple input power with the plasma, resulting in greater dissociation of etch species.

Although a wide variety of high density source types have been developed, they generally fall into one of three categories: electron cyclotron resonance (ECR), helicon resonance, or ICP type sources. All three are in use on production equipment. The main difference is that ECR and helicon sources employ an external magnetic field to shape and contain the plasma, while ICP sources do not [29]. (See Figure 5).

3.5.1. Electron Cyclotron Resonance Plasma

Various methods have been developed for reducing ion energies in the discharge while trying to maintain anisotropic etching. One of the methods involves the addition of magnetic fields configured to reduce electron loss from the plasma and the sample. This method of magnetically enhancing the discharge is called ECR plasma etching. In ECR discharges, free electrons in the plasma are forced to

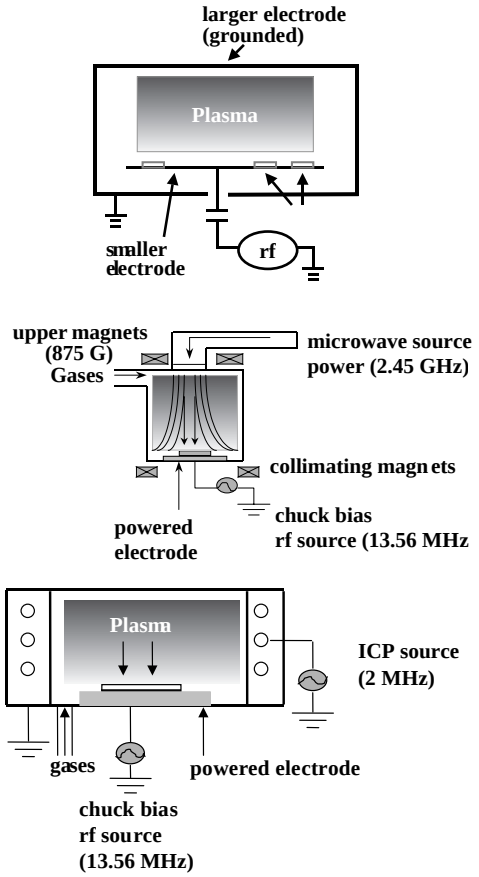


Figure 5. Schematics of typical RIE (top), ECR (center), and ICP (bottom).

orbit about magnetic field lines while absorbing microwave energy. At the cyclotron resonance condition, outer shell electrons from gas molecules in the discharge may also be liberated, leading to a very high degree of ionization in the plasma. Since the motion of the electrons is constrained by the external magnetic field, fewer are lost by collisions with the reactor walls than conventional reactive ion etching (RIE), and therefore the plasma potential relative to ground is much lower. The resultant energies of an ion reaching the sample to be etched are typically $\leq$15 eV. ECR etching should lead to much lower levels of damage than conventional RIE processes. ECR sources operate at the relatively high ion density of 10^{11}–10^{12} cm^{-3} compared with RIE tools ($\sim 10^{9}$ cm^{-3}). ECR tools can also provide independent control of the ion energy and ion flux. Ion energy is controlled by rf or direct current (dc) biasing of the substrate holder while control of the ion and neutral flux is achieved by varying the microwave power and neutral gas pressure. In addition, ECR discharges are capable of low pressure operation due to their efficient dissociation of gases from the discharge.

3.5.2. Inductively Coupled Plasma

Another high density plasma, ICP, has become popular because of disadvantages of the ECR technology. ECR suffers from difficulties in uniformity. The power supply may also limit the scaling of the ECR approach without the development of higher power magnetrons. The primary disadvantages of the ECR technology are the limits due to a commercially available, automatic-tuning microwave power supply, and the physical limits of the magnets required to create a uniform magnetized plasma. ECR tools appear difficult to scale to process wafers larger than 200 mm.

In the ICP geometry, a rf coil encircles the chamber. The important features of this coil are that it carries rf current and generates a magnetic field in the upward and downward directions. The time rate of change of the individual magnetic field generates an electric field. Acceleration of the electrons is determined by the magnitude of the electric field, confining them in a circular motion. The plasma, first formed in the shape of a ring following the path of electrons, will diffuse to the center of the chamber and then downward toward the sample. The electrons, in a circular path, will have only a small chance to be lost to the chamber walls, resulting in low dc self-bias. Ion energy, separated from the ion flux, can be controlled by applying another rf source at the chuck [30]. Unlike ECR plasma sources, there is no resonance between electron motion and the frequency of the driving fields in ICP sources. Therefore, ICP sources have advantages over ECR sources, including easier tuning, scaling up, and lower cost.

The push to continually increase bit densities in magnetic storage devices places emphasis on techniques for patterning submicrometer metallic multilayer structures. The component materials within these multilayers may include NiFe, NiFeCo (both are used for structures based on the GMR effect), NiMnSb (a Heusler alloy potentially useful in advanced spin-valve structures), or the collossal magnetoresistance (CMR) materials $LaCaMnO_3$, $LaSrMnO_3$, and $PrBaCaMnO_3$. A general problem when processing magnetoresistive materials is their relative involatility in conventional dry etching techniques such as RIE.

Practical etch rates may be achieved under high density plasma (HDP) conditions, where the high ion flux is able to promote desorption of halogenated etch products. Both ICP and ECR systems have proven capable of providing the necessary ion-to-reactive-neutral ratio ($>$0.02). We have completed an examination of the etch rates of the three basic classes of materials (i.e., NiFe, NiMnSb, and the perovskite-based CMR materials) in different plasma chemistries and also measured the long-term magnetic and structural stability of dry etched magnetic multilayer stacks. In the latter case, we used MRAM elements as our test vehicle.

The NiFe and NiMnSb layers were deposited on Si (100) substrates by direct current magnetron sputtering from composite targets. Typical layer thicknesses were 5000 Å. Liquid delivery metal organic chemical vapor deposition using 2,2,6,6-tetramethyl-3, 5-hepanedionato (TMHD) precursors [i.e., $La(TMHD)_3$, $Sr(TMHD)_2$, $Mn(TMHD)_3$, and $Ca(TMHD)_2$] was employed to deposit films of $La_{0.41}Ca_{0.59}MnO_3$ on Al_2O_3 (0001) single crystal substrates at 700 °C. The precursors were transported by N_2 carrier gas, with direct injection of O_2 and N_2O as oxidants. Films of $Pr_{0.65}Ba_{0.05}Ca_{0.3}MnO_3$ were deposited on Si (100) in a pulsed laser ablation system (248 nm KeF laser, 10 kHz, 2.5 J/cm^{-2} energy density) at a substrate temperature of 700 °C and

an O_2 partial pressure of 250 mTorr. All of the CMR films were 1500–2500 Å thick.

Finally, the MRAM structures consisted of the following layers deposited on 300 Å of SiN_x on 8500 Å of SiO_2 on Si: 80 Å NiFeCo, 15 Å CoFe, 35 Å Cu, 15 Å CoFe, 80 Å NiFeCo, 200 Å Ta, 550 Å TaN, and 800 Å CrSi. The deposition was performed by Ar ion-assisted sputtering. A 3000-Å-thick SiO_2 mask was patterned by SF_6/Ar RIE as the etch mask for subsequent high density plasma patterning of the metal layers.

The etching was performed in either a Plasma-Therm 790 ICP or a Plasma-Therm SLR 770 ECR reactor. In both systems the samples were thermally bonded to a radio frequency powered (13.56 MHz) chuck which was He-backside cooled. The respective HDP sources were powered up to 1000 W at either 2 MHz (ICP) or 2.45 GHz (ECR). The gases were injected directly into the sources through electronic mass flow controllers at a typical load of 15–20 standard cubic centimeters per minute. We investigated halogen-(Cl_2, BI_3, BBr_3, ICl, IBr, SF_6), CH_4/H_2-, and CO/NH_3-based mixtures since these cover the full range of possible etch products (i.e., metal chlorides, bromides, iodides or fluorides; metal organics or carbonyls).

Magnetic properties before and after plasma etching were determined using superconducting quantum interference device magnetometry (Quantum Design MPMS-5S00) at 4.2 K. Scanning electron microscopy (SEM) was used to examine sidewall smoothness on etched features.

Under RIE conditions (i.e., zero watts HDP source power) we invariably saw net deposition on the samples upon exposure to halogenated mixtures, or essentially no etching with CH_4/H_2 and CO/NH_3 mixtures. Examination of the halogen-plasma exposed surfaces by Auger electron spectroscopy revealed large concentrations of chlorinated residues. Since the halogenated etch products have larger lattice constants than their pure metal constituent and the products are essentially involatile under RIE conditions, one observes a buildup of these species as shown schematically in Figure 6 [31–33].

Table 1 shows a compilation of results for NiFe etching in the different chemistries investigated. The highest rates were achieved with Cl_2/Ar, where the role of the inert gas additive is to provide ion-assisted desorption of the chlorinated etch products. We found that the mass of these inert species also played a role, with Xe providing slightly faster

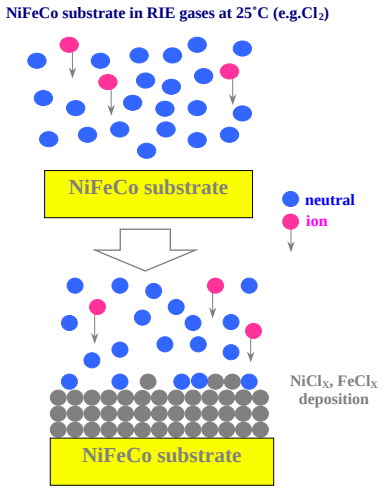


Figure 6. Schematic of involatile layer build-up during Cl_2-based RIE of NiFe at room temperature.

rates than either Ar or He addition. The rates with these Cl_2-based mixtures were approximately a factor of two faster than with pure Ar sputtering. Bromine- or iodine-based plasma chemistries produced lower rates than with chlorine and were close to Ar sputter rates. Both CH_4/H_2/Ar and SF_6/Ar led to extremely low etch rates, while the CO/NH_3 mixture had a slight degree (40%) of chemical enhancement. It has been suggested that the role of the NH_3 is to suppress dissociation of the CO so that carbonyl etch products can form, but an alternative explanation might be that atomic hydrogen scavenges surface carbon species and prevent carbonization of the NiFe surface. The fact that Cl_2-based plasma chemistries produce the fastest rates for NiFe (and plasma chemistries produce the fastest rates for NiFe and NiFeCo) is consistent with the higher vapor pressures of the chlorinated etch products relative to their brominated or iodidated counterparts.

Table 2 shows the corresponding comparisons for NiMnSb. In their cases the Cl_2-based mixtures produce excellent etch rates (1500–5000 Å/min^{-1} for both ICP and ECR tools), but the fastest rates were achieved with SF_6/Ar mixtures. By sharp contrast, NF_3/Ar showed net deposition rather than etching for source powers >100 W or at high NF_3 percentages. The surface under these conditions

Table 1. Comparison of plasma chemistries for NiFe etching.

Chemistry	Typical etch rates	Corrosive	Comments
$10Cl_2/5Ar$	600 Å min^{-1} ICP >1000 Å min^{-1} ECR	Yes	Chemical enhancement of 100% Etch rate with Xe>Ar>He
$13CO/2NH_3$	250 Å min^{-1} ICP 500 Å min^{-1} ECR	No	Chemical enhancement of ~20%–40% CO_2 less effective than CO
$5CH_4/10H_2/5Ar$	<100 Å min^{-1} ICP, ECR	No	Slower than Ar sputtering
$10SF_6/5Ar$	<100 Å min^{-1} ICP, ECR	No	Slower than Ar sputtering
$10BI_3/5Ar$	500 Å min^{-1} ICP, ECR	Yes	Less effective than Cl_2/Ar
$10BBr_3/5Ar$	200 Å min^{-1} ICP, ECR	Yes	Slower than Ar sputtering
10ICl/5AR	500 Å min^{-1} ICP, ECR	Yes	Excellent surface morphology
10IBr/5Ar	500 Å min^{-1} ICP, ECR	Yes	Excellent surface morphology

Table 2. Comparison of plasma chemistries for NiMnSb etching.

Chemistry	Typical etch rates	Corrosive	Comments
$10SF_6/5Ar$	>10,000 Å min^{-1} ICP, ECR	No	Selectivity ≥20 over Al_2O_3
$10NF_3/5Ar$	300 Å min^{-1} ICP, ECR	No	Narrow process window
$10Cl_2/5Ar$	3,000 Å min^{-1} ICP, ECR	Yes	Selectivity ≥5 over Al_2O_3
$10BCl_3/5Ar$	5,000 Å min^{-1} ICP, ECR	Yes	Selectivity ≥5 over Al_2O_3 Attacks native oxide
10ICl/5Ar	1500 Å min^{-1} ICP, ECR	Yes	Threshold ion energy 120 eV
10IBr/5Ar	1500 Å min^{-1} ICP, ECR	Yes	Threshold ion energy 230 eV

showed strong Mn enrichment and were oxidized, with an underlying Sb-deficient region. With all of the plasma chemistries, careful attention had to be paid to the removal of the native oxide prior to the commencement of etching to avoid the presence of a relatively long incubation time.

For the CMR materials, we did not observe any chemical enhancement in etch rate with any of the plasma chemistries discussed (Table 3). The etching was dominated by physical sputtering under all conditions investigated, with etch yields typically <0.1 and relatively high ion energies (>150 eV) needed to initial removal of material.

A key issue with the use of corrosive gas mixtures for etching metallic multilayers is that of postetch stability of the patterned structures. Severe corrosion and delamination of the films is observed in the absence of preventive measures. We examined use of several different postetch treatments. The first was simply rinsing the samples in deionized water immediately upon opening the chamber (which is contained within a N_2 dry box). The samples were then thoroughly dried with filtered N_2. In the other three methods, various *in-situ* plasma cleaning procedures were examined. After Cl_2/Ar etching was complete, the chamber was evacuated for 15 min, and then a 30 mTorr discharge of either H_2, O_2, or SF_6 (500 W source power, 5 W chuck power) was used to clean the residual chlorine for 10 min prior to removal of the samples from the reactor. In these cases, no H_2O rinsing was performed. It should be pointed out that all of these cleaning procedures have been employed previously for removing etch residues after Cl_2-based plasma etching of Al interconnects in Si microelectronics.

Figure 7 shows the magnetization of each of the samples over a period of approximately six months. In each case the samples were simply stored in air between the measurements and no special precautions were taken to prevent corrosion. Each of the cleaning procedures produces samples with extremely stable magnetic characteristics. This is also reflected in their appearance. Figure 8 shows SEM micrographs of patterned MRAM elements three months after Cl_2/Ar etching and postetch cleaning. There is no indication of corrosion on any of the samples and the sidewalls are smooth (to the resolution used in the photos). There is no indication of striations often observed on dry etched features. Note, however, that in the case of O_2 plasma cleaned samples there was a slight decrease in the magnetization per unit volume relative to the samples treated in water or H_2 or SF_6 plasmas. A possible reason for this is that the feature sidewalls become more oxidized than with other treatments, leading to a degradation in magnetic properties.

A comprehensive survey of etching results for magnetic materials in different plasma chemistries has produced the following conclusions:

(i) The optimum chemistry for NiFe is Cl_2/Ar, for NiMnSb is SF_6/Ar, while no chemical enhancement of etch rates for CMR oxides was observed.
(ii) Postetch rinsing in H_2O or *in-situ* plasma cleaning with H_2, O_2, or SF_6 discharges are all effective treatments for removing chlorine etch residues. Of these, only O_2 plasma exposure appears to degrade the magnetic properties of MRAM stacks. Once the residues are removed, there is no change in magnetic or visual properties over a period of ~six months (extent of our study).
(iii) The CO/NH_3 chemistry, while being noncorrosive, produces relatively slow etch rates and is only suitable for patterning of thin (<1000 Å) structures.

Table 4 shows thermochemical data for the potential metal chloride or metal carbonyl etch products for NiFe and

Table 3. Comparison of plasma chemistries for CMR etching.

Chemistry	Typical etch rates	Corrosive	Comments
$10SF_6/Ar$	500 Å min^{-1} $LaCaMnO_3$ ICP, ECR	No	No chemical enhancement
$5CH_4/10H_2/5Ar$	200 Å min^{-1} $LaCaMnO_3$ ICP, ECR	No	No chemical enhancement
$10Cl_2/5Ar$	1500 Å min^{-1} $LaCaMnO_3$ ICP, ECR 900 Å min^{-1} $LaSrMnO_3$ ICP 300 Å min^{-1} $PrBaCaMnO_3$ ICP	Yes	Physically dominated under all conditions for all three materials
$10BI_3/5Ar$	500 Å min^{-1} $LaCaMnO_3$ ICP, ECR	Yes	Etch yield <0.1 threshold ion energy <100 eV
$10BBr_3/5Ar$	500 Å min^{-1} ICP, ECR	Yes	Threshold ion energy ~150 eV

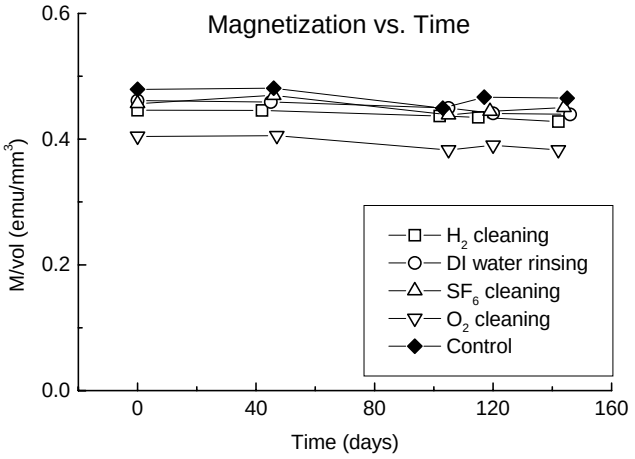


Figure 7. Magnetization of MRAM structures, either unetched or etched in Cl_2/Ar plasmas and subsequently cleaned in water, or in H_2, SF_6 or O_2 plasmas, as a function of subsequent storage time in room ambient.

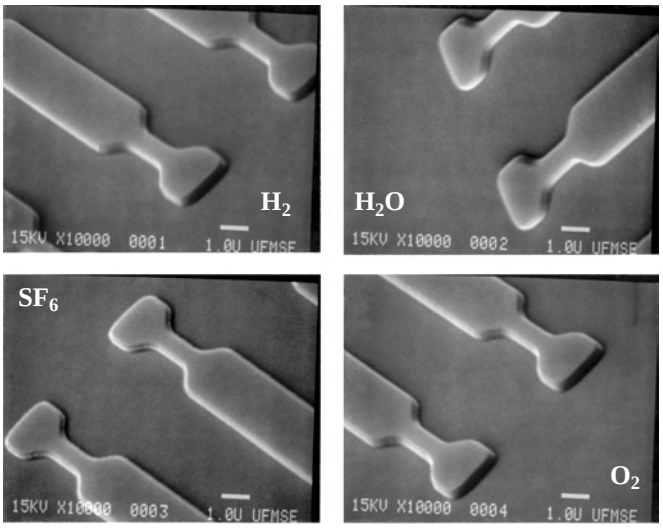


Figure 8. SEM micrographs of MRAM elements after etching in ICP Cl_2/Ar plasmas and subsequent cleaning in H_2, SF_6, or O_2 discharges, or by H_2O rinsing. The micrographs were taken three months after these processes, with the samples having been stored in room ambient.

Table 4. Thermochemical data for potential etch products.

Species	$\Delta H^0_{f.298\ K}$(kJ/mol)	$S^0_{298\ K}$(J/mol K)	$\Delta G^0_{f.298\ K}$(kJ/mol)
$FeCl_2$ (s)	−341.833	117.947	−302.342
$FeCl_3$ (s)	−399.405	142.336	−333.930
$CoCl_2$ (s)	−312.545	109.266	−269.647
$NiCl_2$ (s)	−304.930	98.157	−258.779
$Fe(CO)_5$ (s)	−766.090	337.078	−696.975
$Fe(CO)_5$ (s)	−727.849	439.286	−389.207
$Ni(CO)_4$ (s)	−631.784	319.560	−588.980
$Ni(CO)_4$ (s)	−601.576	415.507	−587.378
$Co(CO)_4$ (s)	−562.100	−337.442	−535.024

NiFeCo in Cl_2 or CO/NH_3 plasmas. From this data we can calculate the Gibbs free energies of reactions of Ni, Fe, and Co with atomic or molecular chlorine and with CO and CO_2. There are several important features of these data in Table 5. First, the reaction of the metals is more favorable with atomic chlorine than with Cl_2, which emphasizes the need for efficient dissociation of the feedstock gas in the plasma source. Second, CO is more reactive with the metals than is CO_2, as we have previously reported in a comparison of the two gases. We emphasize that in a plasma etching environment there will be a strong ion-assisted component to the etch mechanism and the thermodynamic data provide only a guide to the reaction pathways [34–37].

4. DRY ETCHING MECHANISM OF Cu AND MAGNETIC MATERIALS WITH UV ILLUMINATION

In recent years several research groups have studied dry etching of copper for the next generation of metallization in the semiconductor industry, focusing on development of new etch techniques to increase etch rate [38–47]. They used Cl_2 plasmas with or without photon sources using ultraviolet (UV) lasers, UV lamps, illumination, and infrared light. In contrast to conventional dry etching that requires relatively high temperatures (>200 °C) in order to produce practical etch rates, they all reported substantial enhancement of etch rates at low temperatures [48]. Among them Choi and Han first reported high etch rates of about 3000 Å/min at room temperature with Cl_2 discharges in an ICP system.

Magnetic materials such as NiFe and NiFeCo are widely used in sensors, MRAMs, or read/write heads for data storage industry. Due to the relative inertness of these materials there is a strong interest in the development of high density plasma etching processes for them. There are two basic plasma chemistries for the etching of NiFe and NiFeCo under ICP conditions, namely Cl_2 and CO/NH_3. However, the etch rates are still low (<500 Å/min) and are limited by desorption of the etch products such as $NiCl_x$, $FeCl_x$, and $CoCl_x$[49–65]. Cho et al. first reported the effect of UV illumination on the etch rates of the NiFe-based magnetic materials.

Table 5. Thermodynamic data for possible etch reactions.

$Fe(s) + 2Cl(g) \leftrightarrow FeCl_2(s)$	$\Delta G^0_r = -512.9$ (kJ/mol)
$Fe(s) + Cl_2(g) \leftrightarrow FeCl_2(s)$	$\Delta G^0_r = -302.3$ (kJ/mol)
$Fe(s) + 5CO(g) \leftrightarrow Fe(CO)_5$ (1)	$\Delta G^0_r = -11.2$ (kJ/mol)
$\leftrightarrow Fe(CO)_5$ (g)	$\Delta G^0_r = -3.4$ (kJ/mol)
$Fe(s) + 5CO_2(g) \leftrightarrow Fe(CO)_5(1) + 2.5O_2(g)$	$\Delta G^0_r = 1274.9$ (kJ/mol)
$\leftrightarrow Fe(CO)_5(g) + 2.5O_2(g)$	$\Delta G^0_r = 1282.7$ (kJ/mol)
$Ni(s) + 2Cl(g) \leftrightarrow NiCl_2(s)$	$\Delta G^0_r = -359.4$ (kJ/mol)
$Ni(s) + Cl_2(g) \leftrightarrow NiCl_2(s)$	$\Delta G^0_r = -258.8$ (kJ/mol)
$Ni(s) + 4CO(g) \leftrightarrow Ni(CO)_4$ (1)	$\Delta G^0_r = -40.3$ (kJ/mol)
$\leftrightarrow Ni(CO)_4(g)$	$\Delta G^0_r = -38.7$ (kJ/mol)
$Ni(s) + 4CO_2(g) \leftrightarrow Ni(CO)_4(1) + 2O_2(g)$	$\Delta G^0_r = 988.6$ (kJ/mol)
$\leftrightarrow Ni(CO)_4(g) + 2O_2(g)$	$\Delta G^0_r = 990.2$ (kJ/mol)
$Co(s) + 2Cl(g) \leftrightarrow CoCl_2(s)$	$\Delta G^0_r = -380.3$ (kJ/mol)
$Co(s) + Cl_2(g) \leftrightarrow CoCl_2(s)$	$\Delta G^0_r = -269.6$ (kJ/mol)
$Co(s) + 4CO(g) \leftrightarrow Co(CO)_4(g)$	$\Delta G^0_r = 13.6$ (kJ/mol)
$Co(s) + 4CO_2(g) \leftrightarrow Co(CO)_4(1) + 20_2(g)$	$\Delta G^0_r = 1042.5$ (kJ/mol)

Since the etch mechanism with UV illumination has not been studied in detail, in this chapter we propose an etch mechanism of copper and magnetic materials with UV irradiation based on subprocesses occurring in the Cl_2–ICP etching system. We also carried out ICP etching of NiFe and NiFeCo in Cl_2/Ar discharges with or without UV illumination. We found that the chlorination of the copper surface is enhanced with UV irradiation and the absorption of UV photons by metal chlorides is critical to enhance the removal rate of chlorides. The proposed etch mechanism of copper showed good agreement with observed data determined by mass spectrometry, taken from the literature.

There are likely five subprocesses involved in etching of copper with UV illumination:

1. Photodissociation of Cl_2 in gas phase.
2. Surface chlorination.
3. Absorption of UV photons by reaction products.
4. Photoassisted removal of reaction products.
5. Gas-phase reactions between desorbed species.

4.1. Photodissociation of Cl_2

In addition to formation of Cl radicals by electron collision in the bulk plasma, more Cl_2 molecules are dissociated by collision between photons and molecules:

$$Cl_2 + h\nu \rightarrow 2Cl \tag{7}$$

The above reaction readily occurs because the bond strength of Cl_2 is 2.5 eV and the UV photon energies are 2–4 eV. Hence, compared to the plasmas without UV illumination, the photodissociation reaction provides a chlorine-enriched environment, and the reactive chlorine radicals easily take part in surface reactions.

4.2. Surface Chlorination

When the copper surface is exposed to UV radiation, electrons are ejected by the photoelectron effect and are captured by chlorine radicals near the surface, leading to chlorination of the copper surface:

$$Cu(s) + h\nu \rightarrow Cu^+ + e \tag{8}$$

$$xCl(g) + e \rightarrow xCl^-(g) \tag{9}$$

$$Cu(s) + xCl(g) \rightarrow CuCl_x(s) \qquad \Delta G_r^o = -226(\text{kJ/mol}) \tag{10}$$

In contrast to dry etching without UV illumination, UV photons promote the chemistry at the copper surface, resulting in fast deposition of metal chlorides on the surface with low activation energy. This chlorination reaction is induced by photons and the reaction rate is thus a strong function of photon flux or UV intensity. Since the unreacted Cu surface absorbs UV photons and is readily chlorinated by the above reaction, most sites of the copper surface will be chlorinated, leading to the formation of stoichiometric copper chloride.

Under chlorine-enriched conditions, copper chloride may further react with chlorine radicals to form $CuCl_2(s)$

$$CuCl(s) + Cl(g) \rightarrow CuCl_2(s) \qquad \Delta G_r^o = -146\ (\text{kJ/mol}) \tag{11}$$

The chlorination reactions of (10) and (11) may produce the copper chlorides having a Cl content equal to or greater than the stoichiometric ratio (i.e., $x \geq 1$). If there are no UV photons involved, the Cl concentration will be strongly dependent on reaction time. The surface chlorination also weakens Cu–Cu bond strength, resulting in lowering the etch threshold.

4.3. Absorption of UV Radiation by Reaction Products

The enhancement of etch rate with UV illumination indicates that the UV photons play an important role in the etching mechanism. The clue to this suggestion is the optical properties of the reaction products, $CuCl_x$. Tables 6 and 7 show the optical constants of some materials available in terms of reflectivity and adsorption depth. Compared to metal copper, CuCl has a much smaller reflectivity (or longer adsorption depth). This indicates that copper chloride absorbs most UV photon energies, and in turn the photons excite electrons of the reaction products so that the bond strength becomes weaker. This phenomenon may play a key role in increasing the etch rates of copper. If a metal chloride has a low absorption capacity of UV light, the effect of UV illumination will be less significant. Hence, the overall etch process of metals with UV illumination is limited by absorption of UV radiation, which is determined by optical properties of the metal chlorides.

4.4. Photoassisted Removal of Metal Chlorides at the Surface

As soon as the copper chlorides are formed, they absorb UV radiation, resulting in excitation of valence electrons in the $CuCl_x$ layer, weakening bond strengths of $CuCl_x$ and $CuCl_x$–Cu. This could lead to rupture of surface bonds and subsequent desorption. It is noteworthy that CuCl has a direct bandgap of 3.26 eV at 300 K. The UV photon energies in the range 2–4 eV can also be utilized to sublime the copper chloride [1.6 eV for $Cu_3Cl_3(g)$ and 2.2 eV for $CuCl(g)$]. Hence, the possible photoassisted removal process of the chlorinated surface can be described as

$$CuCl(s) + h\nu \rightarrow Cu(g) + Cl(g) \tag{12}$$

$$3CuCl(s) + h\nu \rightarrow Cu_3Cl_3(g) \tag{13}$$

$$CuCl(s) + h\nu \rightarrow CuCl(g) \tag{14}$$

$$Cu(s) + h\nu \rightarrow Cu(g) \tag{15}$$

Table 6. Reflectivity, R, at 298 K (wavelength, λ).

λ (nm)	Cu	CuCl[a]	Ni	Fe	Co	$Ni_{0.8}Fe_{0.2}$[b]	$Ni_{0.8}Fe_{0.13}Co_{0.07}$[b]
300	0.36		0.42	0.35	0.41	0.41	0.41
350	0.40	<0.01	0.44	0.40	0.45	0.43	0.43
400	0.51		0.46	0.48	0.49	0.46	0.46
500	0.59	<0.05	0.54	0.50	0.56	0.52	0.53
600	0.60		0.61	0.52	0.60	0.59	0.60
700	0.96		0.62	0.53	0.62	0.60	0.61

Note: $R = [(n-1)^2 + \kappa^2]/[(n+1)^2 + \kappa^2]$, where n and κ are refractive and absorption indexes at 298 K, respectively, and obtained from Refs. 31 and 33.

[a] From [1].

[b] Estimated using $n = \Sigma x_i n_i$ and $\kappa = \Sigma x_i \kappa_i$, where x_i is atomic fraction of element i.

Table 7. Absorption depth, α^{-1} (nm), at 298 K (wavelength, λ).

λ (nm)	Cu[a]	CuCl[b]	Ni	Fe	Co	$Ni_{0.8}Fe_{0.2}$[c]	$Ni_{0.8}Fe_{0.13}Co_{0.07}$[c]
300	13.9		10.8	13.2	12.1	11.2	11.1
350	14.0	60	12.4	12.6	12.5	12.4	12.3
400	14.4		13.8	12.2	13.0	13.5	13.5
500	15.6	500	14.4	14.2	13.3	14.4	14.3
600	16.2		14.0	16.1	13.8	14.4	14.2
700	13.3		15.5	18.0	14.9	16.0	15.8

Note: Absorption coefficient (nm^{-1}), $\alpha = 4\pi\kappa/\lambda$.
[a] From [33].
[b] From [1].
[c] Estimated using $n = \Sigma x_i n_i$ and $\kappa = \Sigma x_i \kappa_i$, where x_i is atomic fraction of element i.

Photon-sputtering represented by Eq. (15) could occur on the unchlorinated or etched surface due to the lower bond energy of Cu–Cu (1.83 eV)than UV photon energies.

The photon-assisted etch reactions of (12)–(15) are greatly affected by UV intensity because they do not occur without formation of copper chloride and the formation rate is a function of photon flux, implicating a higher etch rate with higher UV intensity.

4.5. Gas-Phase Reactions between Desorbed Species

Some of the desorbed molecules tend to capture electrons, form negative ions, and react with reactive radicals such as Cl and Cu. There are many possible reactions among desorbed species, radicals, and ions in gas phase. They can be summarized as

$$CuCl(g) + e \rightarrow CuCl^{-}(g) \quad (16)$$

$$CuCl_2(g) + e \rightarrow CuCl_2^{-}(g) \quad (17)$$

$$CuCl(g) + Cl(g) \rightarrow CuCl_2(g) \quad (18)$$

$$CuCl^{-}(g) + Cl(g) \rightarrow CuCl_2(g) + e \quad (19)$$

$$CuCl(g) + Cu(g) \rightarrow Cu_2Cl(g) \quad (20)$$

$$CuCl_2^{-}(g) + Cu(g) \rightarrow Cu_2Cl_2(g) + e \quad (21)$$

$$CuCl_2^{-}(g) + Cu(g) + Cl(g) \rightarrow Cu_2Cl_3(g) + e \quad (22)$$

$$CuCl_2^{-}(g) + Cu(g) + Cl(g) \rightarrow Cu_2Cl_3(g) \quad (23)$$

$$Cu_2Cl_2(g) + Cl(g) \rightarrow Cu_2Cl_3(g) \quad (24)$$

$$Cu_2Cl_2(g) + Cu(g) \rightarrow Cu_3Cl_2(g) \quad (25)$$

$$Cu_2Cl_3(g) + Cu(g) \rightarrow Cu_3Cl_3(g) \quad (26)$$

The gas-phase reactions under plasma and UV illumination conditions indicate that the most favored forms of product gas are $CuCl_2$ [Eqs. (17)–(19)] and Cu_2Cl_3 [Eqs. (22)–(24)]. This may be confirmed with examining the relative peak intensities of the observed mass distributions of species using mass spectrometry. Although we may consider photoassisted dissociation of $CuCl(g)$ and $Cu_3Cl_3(g)$, it seems unlikely due to the high bond energy of Cu–Cl, 3.97 eV. The ionization energies of copper chlorides are also quite high: 10.7 eV for $CuCl^{+}$, 9.6 eV for $Cu_2Cl_2^{+}$, and 9.7 eV for $Cu_3Cl_3^{+}$. Hence the ionization of copper chloride gases does not occur under the UV illumination conditions.

Successful etching of NiFe and NiFeCo with Cl_2 plasmas has been reported. However, due to the relative involatility of the etch products such as $NiCl_x$, $FeCl_x$, and $CoCl_x$ the attainable etch rates are quite low ($\leq$500 Å/min). The etch rates of these materials in a high density plasma reactor are function of ion flux, ion energy, and plasma composition. To examine the effect of UV illumination on etch rate we have to know the optical properties of the magnetic materials and reaction products. However, their optical properties are not available yet. There is, however, a substantial increase in etch rate with UV illumination over etching without UV irradiation.

Table 8 summarizes the mass distributions of desorbed gaseous products observed from thermal etching and photon-induced etching of copper. These mass distributions are normalized to the Cu_3Cl_3 intensity. It is seen that for thermal desorption, the major gases are Cu_3Cl_3 at 860 K and CuCl at 920 K, but no $CuCl_2$ and Cu_2Cl_3 are observed.However, UV laser induced etching showed that the dominant gas species were Cu, Cu_2Cl, Cu_2Cl_3, and Cu_3Cl_2 with laser fluence of 0.26 J/cm^2 per pulse with 532 nm, and Cu, Cu_3Cl_2 and Cu_3Cl_3 with 0.66 J/cm^2 and 355 nm. The observation of copper signal is attributed to photon sputtering due to the strong laser energy. By contrast, Kwon et al. reported $CuCl_2$ and Cu_2Cl_3 as the dominant gases observed from ICP etching of copper with UV illumination regardless of chlorine content in copper chlorides. It is also interesting to note that the desorbed gases are Cl, CuCl, $CuCl_2$, Cu_2Cl, Cu_2Cl_2, Cu_2Cl_3, Cu_3Cl_2, and Cu_3Cl_3 in all cases of photon-assisted etching. Furthermore, no observation of $CuCl_2$ and Cu_2Cl_3 during the thermal desorption indicates that their presence in the etching with UV illumination does not originate from cracking of Cu_3Cl_3 by the ionizer of the mass spectrometer. It seems clear that $CuCl_2$ and Cu_2Cl_3 are produced in the ICP etching of copper with UV illumination. Hence, overall these previous results support the etch mechanism proposed in this chapter.

Kwon et al. also confirmed that the copper chloride layer formed with UV illumination had a higher chlorine content than stoichiometry (i.e., $x > 1.0$) regardless of reaction time, and it was composed of $CuCl(s)$ and $CuCl_2(s)$. This is contrary to the much lower chlorine content obtained without

Table 8. Mass distributions of desorbed gases from thermal etching and photon-induced etching of copper at various conditions.

	Mass distribution					
	Thermal etching[28]		UV laser[1]		UV lamp	
Species	860 K	920 K	0.26J/cm^2 532 nm	0.66J/cm^2 355 nm	$CuCl_{0.4}$ 340 nm	$CuCl_{1.2}$ 340 nm
Cl	9	27	28	—	410	350
Cu	20	68	273	47	—	—
CuCl	51	514	89	8	100	100
$CuCl_2$	—	—	121	5	2330	2350
Cu_2Cl	83	82	284	6	130	280
Cu_2Cl_2	10	14	10	13	100	150
Cu_2Cl_3	—	—	282	35	550	2380
Cu_3Cl_2	25	27	141	60	130	270
Cu_3Cl_3	100	100	100	100	100	100

UV illumination. However, this result is attributed to the fact that the copper surface is easily chlorinated under UV illumination because UV photons promote the chemistry at the surface and lower the activation energy. Choi and Han reported the activation energy of 0.12 eV, which is much lower than the energy of 1.6 eV required to sublime $CuCl(s)$ to $Cu_3Cl_3(g)$. This confirms that the dry etching with UV illumination is not a simple thermal desorption, but a non-thermal etch mechanism due to the presence of UV photons.

Figure 9 shows the effect of UV intensity on etch rate and the chlorine content in the copper chloride, adapted from the experiments by Kwon et al. They measured etch rates and Cl contents with varying UV intensity at $Cl_2/N_2 = 1.5$, 2 mTorr, 500 W ICP source power, and room temperature. The etch rates increased linearly with UV intensity up to certain point and then remained almost constant, while the Cl concentration in the copper chloride was independent of UV intensity and maintained at 1.2–1.3, implying coexistence of $CuCl(s)$ and $CuCl_2(s)$. The insensitivity of chlorine atomic ratio to the UV intensity indicates that the UV photon energies used in their experiment are enough for the surface chlorination to occur and to form $CuCl_x$ having $x > 1.0$. The photon-assisted etch of the copper chloride layer occurs very quickly since the $CuCl_x$ layer has low reflectivity but high absorption depth (see Table 2) and absorbs most UV photons as soon as the layer is formed. The increase in etch rate with UV intensity is believed to occur because the photon-assisted removal rate of copper chlorides is faster than the deposition rate of $CuCl_x$, indicating that the deposition rate of copper chloride controls the overall etch process. However, at higher UV intensity the $CuCl_x$ formation rate is also increased due to the increased photon flux, and is thus in equilibrium with the photon-assisted etch rate.

Figure 10 shows the effect of ICP source power in our reactor on etch rates of $Ni_{0.8}Fe_{0.2}$ (top) and $Ni_{0.8}Fe_{0.13}Co_{0.07}$ (bottom) with or without UV illumination in Cl_2 plasmas at 10 sccm Cl_2/5 sccm Ar, 200 W rf chuck power, 2 mTorr, and room temperature. In these experiments an unfiltered 400 W Hg arc lamp was used for UV irradiation. Details of the experiment are described elsewhere. There is net

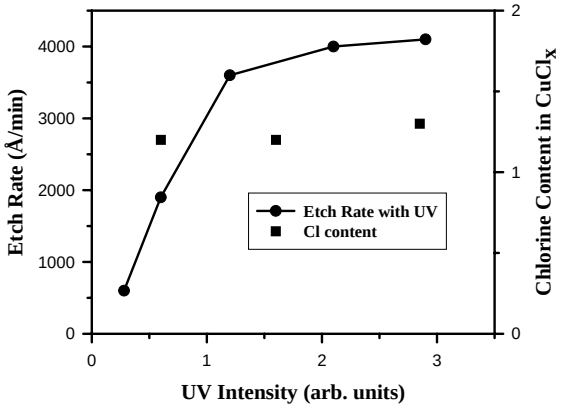


Figure 9. The effect of UV intensity on etch rate of copper and chlorine content in the copper chloride at $Cl_2/N_2 = 1.5$, 2 mTorr, 500 W ICP source power, and room temperature. Adapted from [6], D. E. Heim et al., *IEEE Trans. Magn.* 30, 316 (1994) and [8], R. E. Fontana et al., *IEEE Trans. Magn.* 32, 3440 (1996). © 1994 and 1996, IEEE.

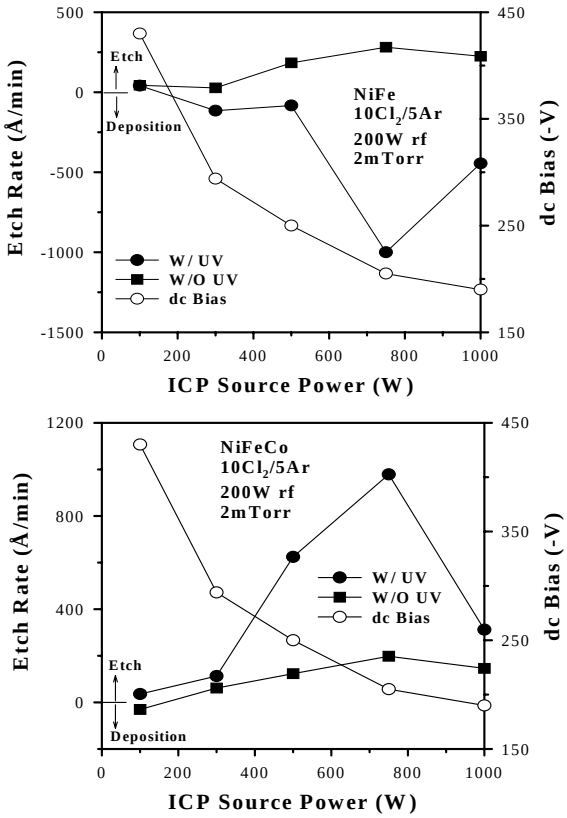


Figure 10. The effect of ICP source power on etch rates of NiFe (top) and NiFeCo (bottom) with or without UV illumination in Cl_2 plasmas at 10 sccm Cl_2/5 sccm Ar, 200 W rf chuck power, 2 mTorr, and room temperature.

deposition observed on NiFe, indicating the rate of formation of metal chlorides is greater than their removal rates. This result also implies that reaction products such as $NiCl_x$ and $FeCl_x$ are not absorbing UV photons. The increase in the formation rate of metal chlorides is attributed to the increased chlorine radicals with UV illumination and the chemistry promoted at the surface by photons. In contrast to NiFe, NiFeCo showed an overall increase in etch rate with UV illumination, especially at moderate ICP source powers (500–800 W). This may be attributed to two factors: (1) lower binding energy of NiFeCo than NiFe (for example, see the Fe–Co phase diagram; addition of Co to Fe lowered the melting point of FeCo alloy) and (2) greater absorption capacity of UV photons by $CoCl_x$ than by $NiCl_x$ and $FeCl_x$. The latter is unlikely because Ni, Fe, and Co are elements all in the same period and the same group so that the alloys and metal chlorides have similar optical properties (Tables 2 and 3), and furthermore the atomic ratio of Co (i.e., 0.07) is too small to affect the overall optical properties of etch products. However, to clearly understand the effect of UV illumination on the dry etching of magnetic materials, more systematic studies, in particular, the dependence of etch rates on optical properties of $NiCl_x$, $FeCl_x$, and $CoCl_x$, need to be conducted.

With the UV-enhanced process, very clearly defined features can be patterned into Cu, as shown in the scanning electron micrographs of Figure 11. The etching was

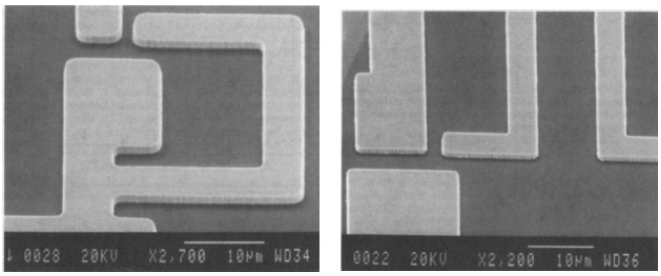


Figure 11. SEM micrographs of features etched into Cu layers on Si, using ICP Cl_2/Ar discharges.

performed at 75 °C in Cl_2/Ar in this case, whereas to achieve similar rates without UV illumination required etch temperatures ≥150 °C.

An etch mechanism with UV illumination was proposed to better understand the ICP etching of copper and magnetic materials. The photodissociation of Cl_2 provides a chlorine-enriched environment near the surface, and UV photons promote the chemistry at the copper surface, leading to fast deposition of metal chlorides on the surface with low activation energy. The proposed model predicts that surface chlorination under UV irradiation produces copper chlorides having Cl content equal to or greater than the stoichiometric ratio (i.e., $x \geq 1$). The overall etch process of metals with UV illumination is limited by absorption of UV radiation, which is determined by optical properties of the metal chlorides. The proposed etch mechanism showed gaseous etch products are CuCl, $CuCl_2$, Cu_2Cl, Cu_2Cl_2, Cu_2Cl_3, Cu_3Cl_2, and Cu_3Cl_3, verified with reported mass spectrometry data [8] and the dominant gas species are $CuCl_2$ and Cu_2Cl_3 in the etching with UV illumination. The Cl_2–ICP etching of magnetic materials with UV illumination showed no enhancement in etch rate for NiFe but a substantial enhancement for NiFeCo mainly due to lower binding energy of NiFeCo. However, to clearly understand the effect of UV illumination on the dry etching of magnetic materials, more systematic studies have to be carried out in terms of the absorption of UV photons by $NiCl_x$, $FeCl_x$, and $CoCl_x$.

5. CONCLUSIONS

Nanoscale patterning of MRAM elements and other magnetic structures is possible using high density plasma etching, combined with appropriate lithography. The choice of plasma chemistry is determined by the materials used in the structure, and both corrosive and noncorrosive chemistries are available.

GLOSSARY

High density plasma A discharge with an average ion density above approximately 10^{11} ions per cubic centimeter.

Ion-induced damage Any disruption of the properties of a material by the energetic ions that bombard its surface during dry etching.

Magnetic random access memory A form of nonvolatile magnetic memory that is promising for data storage.

Plasma chemistry The combination of gases used to create reactive etch products with the sample upon creation of a plasma.

Plasma etching Use of a reactive plasma to etch controlled amounts of a sample exposed to the plasma; also called dry etching.

REFERENCES

1. G. A. Prinz, in "Ultra-Thin Magnetic Structures II" (B. Heinrich and J. A. C. Bland, Eds.). Springer-Verlag, Berlin, 1994.
2. C. H. Tsang, R. E. Fontana, Jr., T. Lin, D. E. Heim, B. A. Gurney, and M. L. Williams, *IBM J. Res. Develop.* 42, 103 (1998).
3. C. H. Tsang, *J. Appl. Phys.* 69, 5393 (1991).
4. R. White, *IEEE Trans. Magn.* 28, 2482 (1992).
5. J. M. Daughton, P. Bade, M. Jenson, and M. Rahmati, *IEEE Trans. Magn.* 28, 2488 (1992).
6. D. E. Heim, R. E. Fontana, Jr., C. H. Tsang, V. Speriosu, B. A. Gurney, and M. L. Williams, *IEEE Trans. Magn.* 30, 316 (1994).
7. M. Parker, K. Coffrey, J. Howard, C. H. Tsang, R. E. Fontana, Jr., and T. Hylton, *IEEE Trans. Magn.* 32, 142 (1996).
8. R. E. Fontana, S. MacDonald, C. H. Tsang, and T. Lin, *IEEE Trans. Magn.* 32, 3440 (1996).
9. B. A. Everitt, A. V. Pohm, and J. M. Daughton, *J. Appl. Phys.* 81, 23639 (1997).
10. S. Wang, F. Liu, K. D. Maranowski, and M. H. Kryder, *IEEE Trans. Magn.* 26, 1689 (1989).
11. S. Wang, E. Louis, F. Wolfson, R. Anderson, and M. H. Kryder, *IEEE Trans. Magn.* 30, 3897 (1994).
12. H. Takano, H. Fukuoka, M. Suzuki, K. Shiiki, and M. Kitadu, *IEEE Trans. Magn.* 27, 4678 (1991).
13. H. Gokan and S. Eho, *J. Vac. Sci. Technol.* 18, 23 (1991).
14. M. J. Vasile and C. J. Mogab, *J. Vac. Sci. Technol. A* 4, 1841 (1986).
15. R. Giridhar, *Jpn. J. Appl. Phys.* 35, 6347 (1996).
16. C. Tsang, M. Chen, T. Yogi, and K. Ju, *IEEE Trans. Magn.* 30, 281 (1994).
17. F. C. M. J. van Delft, *J. Magn. Magn. Mater.* 140–144, 2203 (1995).
18. K. Kinoshita, K. Yamada, and H. Matutera, *IEEE Trans. Magn.* 27, 4888 (1991).
19. M. Balooch, D. S. Fischl, D. R. Olander, and W. J. Siekhaus, *J. Electrochem. Soc.* 135, 2090 (1988).
20. D. S. Fischl and D. W. Hess, *J. Vac. Sci. Technol. B* 6, 1577 (1988).
21. D. W. Hess, *Plasma Chem. Plasma Proc.* 2, 141 (1982).
22. B. Khamsehpour, C. D. W. Wilkinson, and J. N. Chapman, *Appl. Phys. Lett.* 67, 3194 (1995).
23. I. Nakatani, *IEEE Trans. Magn.* 32, 4448 (1996).
24. P. M. Levy, *J. Magn. Magn. Mater.* 140–144, 485 (1995).
25. P. M. Levy, *Solid State Phys.* 47, 367 (1994).
26. See, for example, IBM Storage Division Web site, http://www.storage.ibm.com/hardsoft/diskdrdl/technolo/gmr/gmr.htm.
27. B. El-Kareh, "Fundamentals of Semiconductor Processing Technology." Kluwer Academic, Boston, 1995.
28. D. M. Manos and D. L. Flamm, "Plasma Etching: An Introduction." Academic Press, Boston, 1989.
29. P. Singer, *Semiconductor Int.* 154 (July 1996).
30. B. Gorowitz, R. J. Saia, and E. W. Balch, in "VLSI Electronics Microstructural Science" (N. G. Einspruch, S. S. Cohen, and G. Gildenblat, Eds.), Vol. 15, Chap. 4. Academic Press, Orlando, 1987.
31. A. K. Sinha, H. S. Lindenburger, D. B. Fraser, S. P. Murarka, and E. N. Fuls, *IEEE Trans. Electron Devices.* ED-27, 1425 (1980).
32. T. M. Mayer, J. M. E. Harper, and J. J. Cuomo, *J. Vac. Sci. Technol. A* 3, 1779 (1985).
33. "CRC Handbook of Chemistry and Physics," 72nd ed. CRC Press, Boca Raton, FL, 1989.

34. M. A. Liebermann and A. J. Lichtenburg, "Principles of Plasma Discharges and Materials Processing." Wiley, New York, 1994.
35. K. B. Jung, E. S. Lambers, J. R. Childress, S. J. Pearton, M. Jenson, and A. T. Hurst, Jr., *Appl. Phys. Lett.* 71, 1255 (1997).
36. S. J. Pearton, T. Nakano, and R. A. Gottscho, *J. Appl. Phys.* 69, 4206 (1991).
37. R. J. Shul, M. L. Lovejoy, D. L. Hetherington, D. J. Rieger, J. F. Klem, and M. R. Melloch, *J. Vac. Sci. Technol. B* 13, 27 (1995).
38. R. J. Davis and E. D. Wolf, *J. Vac. Sci. Technol. B* 8, 1798 (1990).
39. G. S. Oehrlein, Y. Zheng, D. Vender, and O. Joubert, *J. Vac. Sci. Technol. A* 12, 323 (1994).
40. J. Mau, *J. Vac. Sci. Technol. B* 6, 652 (1987).
41. D. A. Danner, M. Dalvie, and D. W. Hess, *J. Electrochem. Soc.* 134, 669 (1987).
42. J. W. Lee, J. Hong, and S. J. Pearton, *Appl. Phys. Lett.* 68, 847 (1996).
43. S. J. Pearton, J. W. Lee, E. S. Lambers, J. R. Mileham, C. R. Abernathy, F. Ren, W. S. Hobson, and R. J. Shul, *J. Vac. Sci. Technol. B* 14, 118 (1996).
44. F. Ren, W. S. Hobson, J. R. Lothian, J. Lopata, J. A. Caballero, S. J. Pearton, and M. W. Cole, *Appl. Phys. Lett.* 67, 2497 (1995).
45. J. W. Lee, J. Hong, E. S. Lambers, and S. J. Pearton, *J. Vac. Sci. Technol. B* 15, 652 (1997).
46. K. B. Jung, E. S. Lambers, J. R. Childress, S. J. Pearton, M. Jenson, and A. T. Hurst, Jr., *J. Vac. Sci. Technol. A* 16, 1697 (1998).
47. K. B. Jung, J. Hong, H. Cho, J. R. Childress, S. J. Pearton, M. Jenson, and A. T. Hurst, Jr., *J. Electron. Mater.* 27, 972 (1998).
48. See, for example, "High Density Plasma Sources" (O. A. Popov, Ed.). Noyes, Park Ridge, NJ, 1994.
49. Y. B. Hahn, to be published.
50. B. Vavra, Honeywell SSEC, Plymouth, MN, private communication.
51. See, for example, G. S. Oehrlein and Y. Kurogi, *Mater. Sci. Eng. R* 24, 153 (1998).
52. J. M. Daughton, *Thin Solid Films* 216, 162 (1992).
53. T. Osaka, T. Homma, K. Saito, A. Takekoshi, Y. Yamazuki, and T. Namikawa, *J. Electrochem. Soc.* 139, 1311 (1992).
54. M. Jimbo, K. Komiyama, and S. Tsunashima, *J. Appl. Phys.* 79, 6237 (1996); *J. Magn. Magn. Mater.* 165, 308 (1997).
55. J. A. Caballero, W. J. Geerts, F. Petroff, J.-V. Thiele, D. Weller, and J. R. Childress, *J. Magn. Magn. Mater.* 177–181, 1229 (1998).
56. K. B. Jung, H. Cho, Y. B. Hahn, D. C. Hays, T. Feng, Y. D. Park, J. R. Childress, and S. J. Pearton, *Mater. Sci. Eng. B*, 60, 101 (1999).
57. H. Yoda, H. Iwasaki, T. Kobayashi, A Tsutai, and M. Sahashi, *IEEE Trans. Magn.* 32, 3363 (1996).
58. H. Kanai, K. Yamada, K. Aoshima, Y. Ohtsuku, J. Kane, M. Kanamine, J. Toda, and Y. Mizoshita, *IEEE Trans. Magn.* 32, 3368 (1996).
59. W. P. Jayasekara, S. Wang, and M. H. Kryder, *J. Appl. Phys.* 79, 5880 (1996).
60. C. H. Tsang, T. Lin, S. MacDonald, N. Robertson, S. Santini, M. Duerner, T. Reith, L. Vo, T. Diola, and P. Arnett, *IEEE Trans. Magn.* 33, 2866 (1997).
61. K. Fukuda, M. Sakai, N. Yamanaka, A. Iijima, and M. Matsuzaki, *IEEE Trans. Magn.* 30, 3891 (1994).
62. W. P. Jayasekara, J. Grant, J. A. Bain, A. E. T. Kuiper, and M. H. Kryder, *IEEE Trans. Magn.* 33, 2830 (1997).
63. N. Fukushima, H. Katai, T. Wada, and Y. Horiike, *Japan. J. Appl. Phys.* 35, 2512 (1996).
64. R. E. Lee, *J. Vac. Sci. Technol.* 16, 164 (1979).
65. J. W. Lee, S. J. Pearton, C. J. Santana, J. R. Mileham, E. S. Lambers, C. R. Abernathy, F. Ren, and W. S. Hobson, *J. Electrochem. Soc.* 143, 1093 (1996).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanoscopic Optical Tracers

Wolfgang Schaertl

Universität Mainz, Mainz, Germany

Sabine Schaertl

Evotec OAI AG, Hamburg, Germany

CONTENTS

1. INTRODUCTION

"Nano" signifies one of the most important scientific topics of modern physics, chemistry, and the life sciences. It promises interesting technological developments in the near future, especially in the fields of information technology and modern drug discovery. From a fundamental point of view, nanoscopic probe particles provide the means to explore the structural and dynamic natures of matter on a tiny length scale and improve our understanding of physicochemical processes on a fundamental level.

In this chapter, we will describe the use of specially designed nanoscopic particles in optical tracer experiments. These studies comprise the exploration of transport phenomena in dense interacting colloidal systems, polymer melts, etc. on a micrometer length scale, elucidating for instance interparticle interactions or the mechanism of complex transport processes. A variety of different optical techniques to address these fundamentally important problems has been developed during the last one to two decades, ranging from direct microscopic observation of probe particles [1–5], automatized via modern computers, to more complex fluorescence spectroscopic [6–11] or light-scattering [12–20] techniques. Each technique requires special optical properties of the probe particles, which have to be designed accordingly. For example, nanoscopic probes suitable for fluorescence techniques have to be chemically modified with fluorescent dye molecules or semiconducting metal clusters. On the other hand, the probe particles have to be prepared in such a way that they do not influence the behavior of the surrounding sample itself. These restrictions provide a major challenge for chemical synthesis, which is rewarded by the scientific knowledge one can get from experiments with nanoscopic optical tracers.

A second use of nanoscopic optical tracers lies in modern drug discovery. Potential drugs are identified by their ability to interact with biochemical molecules, which is monitored in an in vitro assay where a preferrably fluorescent probe is used for detection. Here, nanobeads can be used to generate or enhance a signal. Fluorescence confocal single molecule detection [21–27] applied to drug discovery allows the use of highly miniaturized sample volumes and accurate detection of biochemical interactions, and the combination with nanobeads expands the range of applications.

Our contribution is organized as follows: In a first section about optical microrheology, we will start with a short overview of the theoretical background of transport processes in soft condensed matter systems. This part will mainly cover the relationship between standard dynamical-mechanical relaxation experiments using a conventional rheometer, and the information one may get from the optical nanotracer experiments. We will conclude this section with a review of the optical techniques used for optical nanotracing, also providing some illustrative experimental examples taken from the recent literature. The second section of our contribution will deal with the use of nanoparticles in pharmaceutical drug discovery and diagnostics. Here, we will mainly consider recent developments in high throughput drug discovery using fluorescence detection techniques.

2. OPTICAL MICRO- AND NANORHEOLOGY

2.1. Diffusion of Nanoscopic Probe Particles and Mechanical Spectroscopy

Only seven years ago [28], Mason and Weitz set the basis for so-called optical microrheology [28–34]. They have shown that by determination of the single particle trajectories of

ISBN: 1-58883-063-2/$35.00

tiny probe particles embedded in a complex fluid, it is possible to extract the frequency-dependent linear viscoelastic moduli, that is, the flow and elastic properties, of the sample. This new approach provides some important advantages, such as extended frequency range and the ability to probe the dynamic within a micrometer length scale. The latter is especially important for structurally heterogeneous systems, where conventional rheology only can provide a meaningless average result. The motion of a probe particle dispersed in a complex viscoelastic fluid is described by a generalized Langevin equation [35–41]:

$$m\dot{v}(t) = f_R(t) - \int_0^t \zeta(t-\tau)v(\tau)d\tau \tag{1}$$

with m the particle mass and $\dot{v}(t)$ the time derivative of the particle velocity. $f_R(t)$ is the random forces acting on the probe particle and includes both stochastic Brownian forces exerted on the probe particle by thermally incited density fluctuations within the fluid, and, in the case of higher concentrations of probe particles, direct interactions between individual particles. The integral term represents the viscous damping or friction of the fluid and incorporates a generalized time-dependent memory function $\zeta(t)$, which basically includes the viscoelastic moduli of the fluid. For a complex viscoelastic fluid, the fluctuation-dissipation theorem, which describes the relation between the stochastic density fluctuations driving the probe particle motion and the viscous damping of this motion due to energy dissipation (Note: Both effects are caused by the same stochastic Brownian force $f_R(t)$. The so-called fluctuation-dissipation theorem connects diffusive mobility and viscous friction.), is given by

$$\langle f_R(0)f_R(t)\rangle = k_BT\zeta(t) \tag{2a}$$

with $\langle f_R(0)f_R(t)\rangle$ the time-correlation function of the stochastic force and k_BT the thermal energy. For a simple viscous fluid, the memory function $\zeta(t)$ is replaced by the shear viscosity η_0, and Eq. (2a) becomes

$$\langle f_R(0)f_R(t)\rangle = k_BT\eta_0 \tag{2b}$$

Based on Eqs. (1) and (2a), it can be shown that the viscoelastic memory function $\zeta(t)$ is related to the velocity autocorrelation function $\langle v(0)v(t)\rangle$ and hence to the mean squared displacement of the probe particle $\langle \Delta r^2(t)\rangle$ [28]. After Laplace transformation, a generalized Stokes–Einstein relation (GSER) is obtained:

$$\widetilde{G}(s) = s\tilde{\eta}(s) = \frac{k_BT}{\pi as\langle \Delta\tilde{r}^2(s)\rangle} \tag{3}$$

with s the frequency in the Laplace domain, $\widetilde{G}(s)$ the viscoelastic modulus, $\tilde{\eta}(s)$ the frequency-dependent viscosity of the complex fluid, a the radius of the probe particle, and $\langle \Delta\tilde{r}^2(s)\rangle$ its mean squared displacement. For a viscous fluid, the probe particle is freely diffusing, and its mean squared displacement is simply given by

$$\langle \Delta\tilde{r}^2(s)\rangle = 6D/s^2 \tag{4}$$

In this case, the viscosity is independent of frequency s, and Eq. (3) becomes the well-known Stokes–Einstein equation relating self-diffusion coefficient and shear viscosity η_0:

$$\eta_0 = \frac{k_BT}{6\pi aD} \tag{5}$$

In conventional rheological measurements, the complex dynamic-mechanical modulus $G^*(\omega) = G'(\omega) + iG''(\omega)$, with $G'(\omega)$ the elastic or storage modulus representing the elasticity of the sample and $G''(\omega)$ the loss modulus representing the viscous flow properties of the sample, is measured typically with an oscillatory shear rheometer [42]. A thin film of the sample (thickness about 0.1–1 mm, diameter a few cm) is deformed periodically with a motor, and the resulting stress transferred by the sample to a transducer is measured (see Fig. 1).

It is important to note that the stress and, correspondingly, the viscoelastic moduli of the sample usually depend on the deformation frequency ω. To illustrate the significance of the viscoelastic moduli, let us regard the two extreme cases: For an elastic solid body, suchas a rubber ball, which, upon stress, shows neither flow nor permanent deformation, $G'(\omega) \gg 0$ and $G''(\omega) = 0$. On the other hand, for a simple viscous fluid, suchas water or very dilute solutions, $G'(\omega) = 0$ and $G''(\omega) \gg 0$, and, upon stress, all energy is dissipated into the sample. Complex fluids, suchas concentrated colloidal dispersions, polymer gels, etc., have both elastic and viscous components.

To compare $\widetilde{G}(s)$ obtained from microrheological measurement with $G^*(\omega)$ determined by conventional rheology, one could, in principle, calculate first the Laplace transform and then the Fourier transform of $\widetilde{G}(s)$. In their first experiments, Mason and Weitz preferred to fit $\widetilde{G}(s)$ by a functional form and determined $G^*(\omega)$ by analytic continuation of this fitted form. The authors have verified the validity of the sketched formalism by three different experimental examples: (1) A highly concentrated (56 vol%) colloidal suspension of silica particles of radius 210 nm in ethylene glycol studied by diffusing wave spectroscopy (DWS), a special light-scattering technique that will be discussed further beyond; in this case of a concentrated colloidal dispersion studied by DWS, probe particles and sample are identical: The viscoelastic behavior is mainly determined by direct interactions between the nanoscopic colloidal particles. (2) An emulsion of uniformly sized oil droplets of radius 530 nm, stabilized with a surfactant and suspended in water, volume fraction of the droplets 62%; here, as in

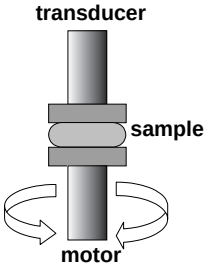


Figure 1. Sketch of a typical oscillatory shear rheometer.

case of example (1), probe particles and sample are identical. (3) An entangled polymer solution of high molecular weight polyethylene oxide dissolved in water, containing a small amount (2 vol%) of nano-sized polystyrene latex spheres of radius 210 nm as optical probe particles; in this last example of a polymer gel, the latex probe particles only serve to visualize the viscoelastic behavior of the polymer network but do not influence it. In all three cases, Mason and Weitz [28] found excellent agreement between the viscoelastic moduli determined by standard rheological measurements and by the new microrheological approach using optical tracer techniques.

Here, it is important to mention that the generalized Stokes–Einstein equation (Eq. (3)), which relates the single probe particle dynamics to the complex viscoelastic modulus, is only valid at certain length scales and within a limited frequency regime [43–47]: If the medium is inhomogeneous on the particle length scale, the tracer particle will follow the easy way, and its dynamics will predominantly sample local regions that provide the least resistance to deformation. Therefore, these soft cavities are oversampled. Only if the size of the probe particle is much larger than the characteristic length scale of the viscoelastic fluid, the single-particle dynamics reflect the overall viscoelasticity of the dispersing medium. For this reason, optical probes for microrheological measurements are usually several hundred nanometers in size. A recent experimental development provided a solution to this problem: By simultaneous determination of the correlated displacement of two probe particles [48, 49] separated by a distance that is large compared both to the particle size and the characteristic length scale of the viscoelastic medium, using for example optical video microscopy, the result becomes independent of both size and shape of the probe particles. This correlated two-particle displacement, also called two-point microrheology, directly reflects the bulk viscoelasticity of the medium and does not suffer from the length scale restrictions present in studies of single-particle dynamics in microscopically heterogeneous media.

The GSER (Eq. (3)) is not only limited to probe particles larger than the characteristic length scale of a heterogeneous viscoelastic medium, for example a polymer network, but is also restricted in respect to the frequency window $\omega_L < \omega < \omega_U$: The lower frequency ω_L is given by the longitudinal compression mode of the polymer network, and the upper frequency cutoff originates from fluid inertia effects. Due to the important perspectives for the experimentalist, the validity of the GSER has recently become a very important subject in the field of theoretical physics. For a more detailed discussion, which is far beyond the scope of this contribution, we therefore refer to the literature [35, 36, 40, 41, 44–47].

We have tried to sketch how the microscopic dynamics of an optical probe, also called microrheology, and conventional rheological techniques are connected. This formalism, developed just a few years ago, opens up a rich field for the experimentalist: Now, not only can the dynamics of optical probes be used to study simple viscous fluids where the Stokes–Einstein equation holds (Eqs. (4) and (5)), but it is also possible to investigate complex viscoelastic samples and even to determine the viscoelastic moduli in a frequency regime not accessible by conventional rheological methods.

2.2. Experimental Techniques

In this section, we will try to illustrate the importance of optical microrheology by a presentation of various recently developed experimental methods. The interested reader is also referred to a short but comprehensive review about optical microrheology techniques by Mukhopadhyay and Granick [50] that has just been published. All these techniques have in common the fact that their fundamental experimental quantity is the mean squared displacement of a nano- to micron-sized probe particle determined by optical methods. However, they differ strongly in time and length resolution, as well as in the optical properties of their probe particles. Depending on the scientific problem, the suitable technique or even a combination of various methods has to be chosen carefully.

2.2.1. Optical Video Microscopy

In optical video microscopy [1–5], the motion of a nanoscopic probe particle is directly observed with the help of an optical microscope. Images of the particle positions are recorded with a video camera and fed to a personal computer. The major advantage of this technique is that it is capable of tracking many probe particles simultaneously. For this reason, it also allows one to measure the correlated motion of two individual particles necessary for the 2-point microrheology [49] mentioned above. Another advantage is the very small sample volume necessary for microscopic studies ($<100\ \mu$L). However, serious disadvantages of optical videomicroscopy are its limited spatial resolution (>100 nm) and, more crucially, the limited time resolution of modern video hardware (ca. 0.001–0.01 seconds). Therefore, the technique is not suited to study very elastic fluids, where rapid particle motion on a very short length scale is to be expected. In general, currently only the long-time or, correspondingly, the low-frequency regime of the dynamic mechanical modulus $G^*(\omega)$ of a viscoelastic fluid can be studied by videomicroscopy. In the near future, new technical developments, for example high speed cameras and faster computers, will enhance the time scale of this technique to much shorter times (higher frequencies).

Standard optical microscopic techniques are limited to very thin samples, restricting the observable particle motion to two dimensions. Therefore, well-suited samples are biological monolayers or bilayers, such as cell membranes. Here, optical microscopy, where already one single particle can be detected, is much better suited than light scattering techniques, which, due to the low contrast of an individual particle, need to average the signal from many probes, and therefore are only suited for bulky samples. The restriction of optical microscopy to two dimensions can be overcome by confocal microscopy [51–55], which, although on a very slow time scale of seconds, allows one to change the plane of focus systematically within a sample, thereby creating a three-dimensional microsopic image.

For videomicroscopy, it is necessary to distinguish clearly the probe particles from the surrounding medium by sufficient optical contrast. This optical contrast of the

probe particles may be based on fluorescence (fluorescence microscopy), for which purpose the probes have to be labeled with fluorescent dyes or even semiconducting metal clusters. The latter are much better suited because of their higher stability in case of irradiation [56–59]. It is also possible to tune the wavelength of both excitation and fluorescence via the size of the metal clusters [60, 61]. Alternatively, dark field microscopy [62, 63] may be applied, where the probe particles are distinguished from the surrounding medium by their enhanced light scattering power. Compared to conventional optical transmission microscopy, both the fluorescence and the dark field techniques have the advantage that tracers even much smaller than the wavelength of light (400–800 nm) are clearly visible and therefore can be used as probe particles.

Next, let us consider a recent example of single particle tracking by optical microscopy: The diffusion of individual fluorescent tracer particles has been studied in thin (d = 2–3 μm) concentrated colloidal suspension layers [5]. For this purpose, a few fluorescent tracers (<1 vol%) of radius 260 nm have been added to a dispersion of nonfluorescent colloidal particles of the same size suspended in water. The preparation of suitable tracers has been quite simple: To an aqueous dispersion of the colloidal particles, consisting of the two monomers acrylamide and styrene and prepared by standard radical emulsion polymerization, a solution of the hydrophobic strongly fluorescent dye Nile Red [64–66] in the organic solvent chloroform has been added dropwise. The latex particles are swollen by the organic solution, and, after evaporation of the chloroform, the fluorescent dye, which is not soluble in water, remains within the particles. It should be stressed that the fluorescent label in this case is only incorporated and not chemically attached to the nanoscopic probe particle. A few drops of this tracer dispersion are added to a concentrated dispersion of nonfluorescent particles. Under a fluorescence microscope, only the dye-labeled tracer particles are visible as sketched in Figure 2.

Since the fluorescent tracer particles can be distinguished individually, their single particle dynamics can be studied. The shortest time scale of the measurements of about 0.1 seconds was limited by the digital frame grabber, the interface translating the signal from a video camera on top of the fluorescence microscope into a digital picture that can be analyzed by computer. The smallest length scale of the experiment as well as the length scale resolution is about 250 nm, that is, half the wavelength of visible light. The position of the center of an individual tracer particle is determined by a special computer algorithm and is stored as a function of time. From these raw data, the mean squared displacement $\langle \Delta r^2(t) \rangle$ can be calculated, which yields the self-diffusion coefficient D (see Eq. (4)).

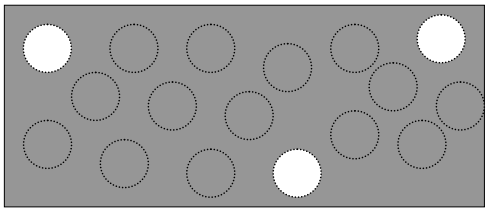


Figure 2. Sketch of the image obtained by fluorescence microscopy from a colloidal dispersion containing nonfluorescent (gray) and fluorescent (bright) colloidal nanoparticles.

In this example, all particles influence the dynamic behavior of the system, but only the tracers are observed. Their individual motion reflects interactions with other invisible colloidal nanoparticles. From the dependence of single particle motion on total particle concentration, the range of the interparticle interactions can be deduced. In case of hard spherical particles, for instance, the diffusion coefficient is expected to depend on particle volume fraction φ roughly as [67–71]:

$$D(\varphi) = D_0(1 - 2\varphi) \tag{6}$$

with D_0 the diffusion coefficient found for very dilute colloidal dispersions and given by the Stokes–Einstein equation (Eq. (5)). The acrylamide-styrene latex particles in aqueous dispersion have been found to show the hard sphere behavior described by Eq. (6), although they contained a small number of surface charges originating from the polymerization initiator potassium peroxodisulfate. This example illustrates the possibility to determine the range of interparticle interactions from the probe dynamics of individual optical nanotracers.

The tracers described above have been prepared in a very primitive way by simply incorporating fluorescent dye molecules by solution swelling. A more sophisticated way is the preparation of cross-linked colloidal nanoparticles with dye molecules chemically incorporated in their network structure [72, 73]. To avoid labels sticking out from the particle surface and thereby changing the interparticle interactions of the probe particles, a protective shell that carries no labels can be formed around the fluorescent core. These particles can be isolated and redispersed in various organic solvents while maintaining their topology.

As a second important example for the application of optical video microscopy, we next will consider in more detail the two-point microrheology mentioned above: This technique is very important to determine the macroscopic dynamic mechanical properties of samples that are heterogeneous on a microscopic length scale. As presented in the theory section, conventional microrheology is based on the generalized Stokes–Einstein relation (Eq. (3)). This approach does not agree with conventional rheological methods in case of inhomogeneous samples. To solve this problem, Crocker et al. [49] have used multiparticle video tracking to measure the displacement vectors of several tracers simultaneously. Their experimental quantity of interest is the ensemble averaged tensor product of the tracer displacements:

$$D_{\alpha\beta}(r, \tau) = \langle \Delta r_\alpha^i(t, \tau) \Delta r_\beta^j(t, \tau) \delta[r - R^{ij}(t)] \rangle_{i \neq j,\, t} \tag{7}$$

where i and j label different particles, α and β label different coordinates, and R^{ij} is the distance between particles i and j. δ symbolizes the delta function and becomes 0 if $r \neq R^{ij}(t)$. In the limit of long distances $r \gg a$, one gets:

$$\widetilde{D}_\pi(r, s) = \frac{k_B T}{2\pi r s \widetilde{G}(s)} \tag{8}$$

where $\widetilde{D}_\pi(r, s)$ is the Laplace transform of $D_\pi(r, \tau)$ and the off-diagonal tensor elements vanish. Here, it should be noted that Eq. (8) in contrast to Eq. (3) contains no dependence on probe particle size a. Therefore, the experimental quantity $D_{\alpha\beta}(r, \tau)$ is independent of the size and shape of the probe particles. Based on Eqs. (7) and (8), a distinct mean squared displacement can be defined:

$$\langle \Delta r^2(\tau) \rangle_D = \frac{2r}{a} D_\pi(r, \tau) \tag{9}$$

Crocker et al. [49] use this distinct mean squared displacement instead of the single particle displacement to calculate the dynamic mechanical modulus from the GSER (Eq. (3)). Whereas the Brownian motion of a single tracer is driven by the superposition of dynamical modes with wavelengths greater than the particle size a, the correlated motion of two tracers is driven by those modes with wavelengths greater than their separation r rather than a. Thus, the correlated motion of two tracers separated by more than the coarse-graining length scale in an inhomogeneous medium will depend only on the coarse-grained $G^*(\omega)$, even if the size of the tracers themselves is much smaller than the characteristic length scale of the sample.

As an experimental control, Crocker et al. [49] studied dispersions of submicron fluorescent beads as tracers ($a =$ 235 nm) dispersed in a glycerol/water mixture with epifluorescence. A few hundred particles were imaged simultaneously, and a series of images of total duration 10 minutes was recorded on a videotape, which later could be analyzed by computer. As expected for a homogeneous medium, they found good agreement between single particle mean squared displacement and distinct mean squared displacement. To demonstrate the effectiveness of the new approach for the exploration of inhomogeneous fluids, Crocker et al. [49] compared two-point microrheological measurements of an aqueous guar solution containing high molecular weight mesoscopic aggregates with those from a mechanical rheometer. Dynamic mechanical moduli determined from the distinct displacements agreed very well with the rheological results, whereas the moduli determined from the single particle displacements were far too small and also showed the wrong frequency dependence (since a small single tracer particle oversamples the "softer" regions of a heterogeneous fluid; see above). The heterogeneous guar system therefore nicely demonstrated the superiority of two-point microrheology in case of microscopically heterogeneous samples. As an additional application, the authors studied the dynamics of fluorescent tracers within a filament network of polymerized F-actin obtained from purified rabbit muscle.

2.2.2. Dynamic Light Scattering Using Tracers

Standard Dynamic Light Scattering Dynamic light scattering (DLS) has become a common tool for particle sizing [74–77] by studying the dynamics of the particles of interest in a very dilute dispersion. More recently, DLS also has been applied to study diffusion in highly concentrated colloidal dispersions to explore their phase behavior and the interparticle interactions [15, 16, 19, 78, 79]. The typical setup of a DLS experiment is sketched in Figure 3.

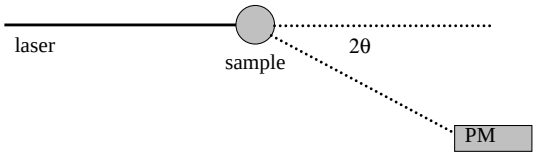


Figure 3. Setup for DLS experiments (top view).

A coherent laser beam is focused onto the sample cell, typically a cylindrical cuvette of diameter 1 cm made from Suprasil glass. The time-dependent scattered light intensity $I(t)$ is detected under an observation angle 2θ with a photomultiplier. It is important that the sample be nearly optical transparent since multiple scattering spoils the simple relation between observation scattering angle 2θ and the inverse length scale of the experiment q (see Eq. (10)). However, recently a new light-scattering technique called diffusing wave spectroscopy (DWS) [80–86] has been developed, which is explicitly based on multiple scattering events and therefore is suitable to investigate tracer particle dynamics also in highly concentrated turbid samples. This method will be reviewed in the next section.

Let us consider the standard DLS technique in some more detail: Due to the Brownian motion of the scattering particles, the intensity pattern $I(t)$ shows characteristic fluctuations with time. For data evaluation, the time intensity autocorrelation function is determined using a hardware correlator. In the case of simple Brownian motion in a viscous fluid medium, this correlation function provides a direct measure for the self-diffusion coefficient D

$$\begin{aligned} \langle I(t)I(t+\tau) \rangle_t &= g_2(q, \tau) = 1 + e^{-2Dq^2\tau}, \\ q &= \frac{4\pi n_D \sin \vartheta}{\lambda} \end{aligned} \tag{10}$$

with τ the correlation time, q the scattering vector, n_D the refractive index of the medium, and λ the wavelength of the laser light. To investigate more complex dynamics of scattering probe particles, it is possible to calculate the mean squared displacement $\langle \Delta r^2(t) \rangle$ from $g_2(q, \tau)$. Therefore, DLS experiments measure the same quantity as determined by direct observation of the probe particles with an optical microscope ($\langle \Delta r^2(t) \rangle$), at comparable length scales of 100 nm–10 μm but on much shorter time scales (as short as 100 nanoseconds).

In case of samples containing a large amount of scattering particles, the DLS measurement no longer probes single particle dynamics, but collective diffusion, which is based on multiple-particle density fluctuations. To measure single particle diffusion in such a multiple-particle system, analogous to video microscopy a special host-tracer technique has to be applied: Matrix particles are dispersed in a refractive index matching solvent to minimize their scattering intensity. A small amount of strongly scattering tracer particles is added to the dispersion. Therefore, basically the situation depicted in Figure 2 is obtained: only part of the particles present in the sample is visible to the experiment. It should be noted here that the scattering intensity strongly depends on both particle size a and the refractive index difference of the probe particles in respect to the surrounding medium Δn_D.

For particles with radius $a < 100$ nm, the scattered intensity per particle in principle is given by

$$\langle I(q,t)\rangle \sim a^3 \Delta n_D^2 \tag{11}$$

Using an adequate host-tracer system, DLS experiments are sensitive only to the tracer particles, the sample therefore optically corresponding to a very dilute dispersion where single particle mobility easily can be determined (in an invisible but highly concentrated matrix). This approach has been used previously to study self-diffusion in colloidal dispersions in the glass transition regime [18, 87–91], that is, at very high particle concentrations close to the liquid–solid transition. Both host and tracer particles have been quite large ($a \sim 100$ nm). Therefore, a comparatively small difference in refractive index of host and tracer particles ($\Delta n_D \sim 0.10$–0.15) has been sufficient.

Also, tracer light scattering experiments on much smaller colloidal particles ($a \sim 10$–20 nm) have been reported [92, 93]. To obtain a sufficient scattering contrast between host and tracer particles, a much larger value of Δn_D is needed in this case. To achieve this goal, the technique of resonance enhanced Rayleigh scattering reported by Stanton et al. [94] can be applied: This method is based on choosing the wavelength of the incident laser light close to an absorption band of the scattering particles, which leads to an extreme enhancement of the scattered intensity. As adequate tracer particles, highly cross-linked polyorganosiloxane μ-gels containing small gold clusters [95], which absorb light in the wavelength regime of the incident laser beam (514 nm), have been prepared. Liz-Marzán and Philipse [93] previously have employed a similar system, silica-coated gold clusters, in their tracer DLS measurements of concentrated colloidal suspensions.

Using light absorbing nanoscopic tracers in DLS experiments, an unusual and interesting phenomenon has been encountered: Whereas intensity time correlation functions usually are expected to show a simple exponential decay, from which the self-diffusion coefficient can be determined (see Eq. (10)), in this case strongly oscillating signals have been found [92, 96]. This phenomenon has been attributed to convective flow of the scattering probe particles resulting from the heating of the illuminated part of the sample by the laser beam. Based on a detailed data analysis (see Eqs. (12a) and 12(b)), it has been possible for the first time to quantify a convective flow pattern using dynamic light scattering with optical nanotracers [92]. The tracer particles exhibit a superposition of their ordinary diffusive Brownian motion and a ballistic part following the convective flux of the dispersion, yielding the following signal in a DLS experiment:

$$\begin{aligned} g_1(\vec{q},\tau) &= [g_2(\vec{q},\tau) - 1]^{0.5} \\ &= \{\exp(i\vec{q}\vec{v}\tau) + \exp[i\vec{q}(-\vec{v})\tau]\}\exp(-Dq^2\tau) \\ &= \ldots\ldots\ldots = 2\cos(\vec{q}\vec{v}\tau)\exp(-Dq^2\tau) \end{aligned} \tag{12a}$$

$$\vec{q}\vec{v} = \frac{4\pi n_D \sin(\vartheta)}{\lambda}\cos(\vartheta)v = \frac{2\pi n_D}{\lambda}\sin(2\vartheta)v \tag{12b}$$

In Eq. (12a), the sum of two complex exponentials corresponds to the convective motion of the scattering nanotracers, whereas the simple exponential term contains the diffusive part. As a consequence, the correlation function becomes a superposition of an oscillatory part and an exponential decay, and basically looks like a damped oscillation function. The convective flow pattern corresponding to Eq. (12a) is sketched in Figure 4.

In case of this simple convection geometry, the sum of two complex exponentials simply yields a cosine function as shown in Eq. (12a). The argument of this cosine function, and therefore the frequency of the oscillations found in the time intensity correlation function, $g_1(\vec{q},\tau)$ depends both on scattering vector q and particle flow velocity v (see Eq. (12b)).

This example illustrates how optical nanotracers can be used also to investigate complex transport phenomena suchas thermal convection. Here, we should stress once more that DLS experiments are limited to optically transparent samples with only a small amount of scattering probe particles present in the illuminated sample volume. However, many viscoelastic samples of interest are turbid and show strong multiple scattering. A solution to this problem is provided by diffusing wave spectroscopy (DWS) [31, 80–86], a revolutionary dynamic light scattering technique presented in the next section.

Diffusing Wave Spectroscopy Whereas in standard DLS experiments a single scattering event from a moving probe particle is detected, in DWS basically the scattered light itself is diffusing: Many scattering probe particles participate in a typical light path through the medium, and very small displacements of these particles, due to their diffusive Brownian motion, cause a decorrelation of the intensity of scattered light. As a consequence, DWS measurements are sensitive to probe particle displacements on length scales much smaller than the wavelength of light, that is several Å. In addition, these displacements may be probed over the extraordinary wide range of time scales from 10^{-8} to 10^5 seconds. Microrheological measurements using DWS have recently been applied to a variety of systems, ranging from turbid colloidal suspensions in which the scattering particles are intrinsic to the medium [97, 98], to weakly scattering media, such as polymer solutions and gels, rendered turbid by the addition of colloidal probe particles [30, 97, 99–103].

The DWS method is based on the same principle as standard DLS: the time-averaged intensity autocorrelation function is determined by collecting the scattered intensity $I(t)$

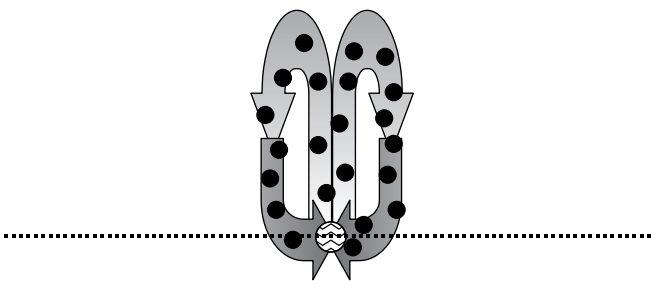


Figure 4. Convective flow pattern of a dispersion of light absorbing nanotracers (●) heated by the incident laser beam (≋) during a DLS experiment. Only flow components within the scattering plane (dotted line) contribute to the correlation function (see Eq. (12a)).

from a single speckle (= patch of scattered light, fluctuating in intensity with time) over a sufficiently long period of time and calculating $g_2(\tau)$ with a hardware correlator. Next, the time-averaged field autocorrelation function $g_1(\tau)$ is calculated. As in case of DLS, the mean squared displacement of a single probe particle $\langle \Delta r^2(t) \rangle$ can be determined from $g_1(\tau)$. The main difference between DLS and DWS is that the scattering vector q plays no role in DWS experiments, that is, the length scale is not defined by the wavelength of the laser and the scattering angle as in DLS.

The importance of DWS from the point of view of microrheological measurements, besides its general applicability to turbid samples, is based on two technical improvements of the standard DWS technique [104–106]. Recently, a new procedure has been developed to extend DWS measurements to the study of non-ergodic media, that is, complex viscoelastic fluids with dynamic heterogeneities. In such systems, the motions of the embedded scattering particles are hindered by spatially dependent elastic forces. Hence, each particle only probes a restricted portion of its total phase space, which, due to the disordered nature of the medium, varies from particle to particle. As a result, each speckle contains a distinct "frozen" portion of the scattered intensity in addition to a fluctuating part, and the temporal intensity profile of each speckle is different. For such non-ergodic materials, an explicit ensemble average must be performed in order to obtain a meaningful autocorrelation function. As Scheffold et al. have shown [106], this ensemble average can be performed by simply placing an ergodic, turbid "reference" sample cell in series after the primary non-ergodic sample of interest. As a suitable reference, a turbid suspension of strongly scattering colloidal nanoparticles has been used. The thermal fluctuations of this ergodic reference cell spatially average the non-ergodic signal from the sample cell. Principal requirements of the reference cell are that it be less turbid than the sample cell and that its characteristic decorrelation time be much longer than the longest correlation time to be measured in the non-ergodic sample of interest. The longest time scale of this two-cell DWS technique [106] is therefore limited by the fluctuation time of the reference cell, which is typically in the order of seconds for colloidal particles in a viscous Newtonian fluid. One drawback of this technique is that, as in conventional DLS, the correlation function is calculated by averaging the time-dependent scattered intensity with respect to a reference time over the duration of the experiment. To obtain sufficient statistical data accuracy, the total measurement time therefore has to be much longer than the longest correlation time of $g_2(\tau)$. If the scattering properties or the dynamical properties of the sample change during the duration of the experiment, the correlation function becomes meaningless. Thus, like standard DLS, the two-cell DWS technique is not useful for systems that are evolving on experimental time scales.

To shorten the experimental time scale significantly, another technique to detect the scattered intensities of multiple speckles simultaneously and partially replace the time average in the intensity correlation function by an ensemble average over various speckles recently has been developed: multispeckle correlation spectroscopy using a charge coupled device (CCD) camera as an area detector [107–109]. The CCD chip here basically works like an array of photomultipliers, and several hundred speckles can be monitored simultaneously. As a consequence, the total duration of the experiment can be shortened by a factor of more than 100 compared to the single-channel method using only time averaging for the correlation function. The shortest time scale accessible by the multispeckle CCD camera based approach is limited by the frame rate of typical CCD systems, which nowadays typically is in the order of 1,000 Hz.

Correlation times covered by two-cell DWS range from 10^{-8} to 10^{0} seconds. On the other hand, the time regime of the CCD multispeckle approach ranges from 10^{-3} to 10^{5} seconds. Used in concert, these two techniques can probe dynamics in complex viscoelastic samples, even if those are turbid and dynamically heterogeneous (non-ergodic), on the extraordinary wide range of timescales from 10^{-8} to 10^{5} seconds, at a length scale of probe particle displacements of several Å.

2.2.3. Forced Rayleigh Scattering

In the preceding section, we have seen that dynamic light scattering techniques, due to the time averaging necessary for calculating the intensity correlation function with sufficient statistical accuracy, afford a total measurement duration much longer than the longest correlation time. Therefore, their correlation time scale is limited to typically 10^{0} seconds. Using the multiple speckle approach with a CCD area detector [107–109], the time average can be partially replaced by an ensemble average, which shortens the measurement time significantly and provides access to a correlation time scale up to 10^{5} seconds. To measure the ultraslow dynamics of nanoscopic probe particles encountered, for example, in polymer melts, the presented correlation techniques are fairly inadequate. Here, a holographic grating technique, forced Rayleigh scattering (FRS) [110–112], provides the ultimate solution. This optical tracer technique practically holds the world record in the study of very slow diffusional processes: It covers diffusion coefficients in the range from 10^{-10} down to 10^{-20} m^2s^{-1}. Standard DLS, for comparison, typically covers diffusion coefficients in the range 10^{-5} to 10^{-14} m^2s^{-1}. By combination of standard tracer DLS with the multispeckle approach, this regime can be enhanced down to 10^{-17} m^2s^{-1}.

Why is FRS so efficient in exploring ultraslow dynamical processes? The typical experimental setup of a FRS experiment is sketched in Figure 5.

FRS is based on the following principle [110, 112]: The interference of two coherent laser beams (I, II) creates a holographic grating by selectively destroying photoreactive dye labels within the sample. These labels are chemically bound or physically incorporated into the nanoscopic probe particles whose dynamics are to be studied. The holographic grating is a periodic arrangement of bleached and nonbleached particles and therefore corresponds to a periodic spatial modulation of the optical properties of the sample, that is, refractive index and light absorption coefficient. Thermally incited diffusion of the labeled particles (Brownian motion, see above) leads to a successive destruction of the ordered holographic lattice with time by mixing probe particles containing bleached and nonbleached labels.

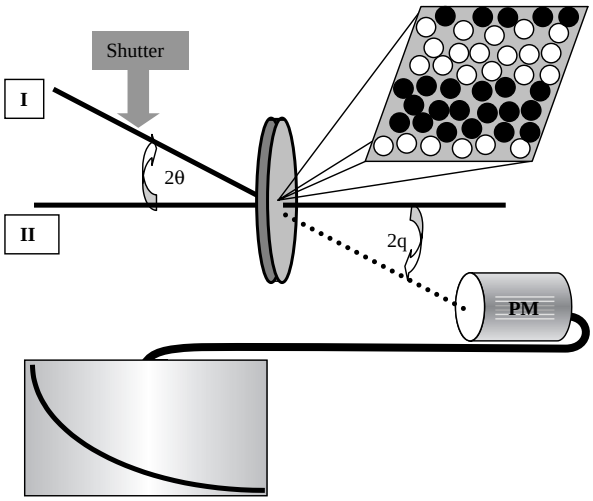


Figure 5. Experimental setup of an FRS experiment.

The actual status of the holographic lattice and thereby the progress of its destruction can be monitored by Bragg diffraction of an attenuated laser beam at the lattice. For nanoscopic optical tracers in a viscous medium, the Bragg diffracted intensity as a function of time is simply given by

$$I(q,t) = A\exp(-2Dq^2t) \tag{13}$$

The scattering vector q is defined by the lattice constant of the holographic lattice, which itself is determined by the interference angle between the two laser beams according to

$$q = \frac{2\pi}{d} = \frac{4\pi\sin(\theta)}{\lambda} \tag{14}$$

Eq. (13) looks identical to Eq. (10), which was used to extract the self-diffusion coefficient from DLS experiments. Note, however, that in DLS the measured scattered intensity shows a fluctuating pattern, and the exponentially decaying function $g_2(q,\tau)$ is a time-averaged intensity correlation function, whereas in FRS experiments the measured scattered intensity itself exhibits the exponential decay. Therefore, in case of the FRS experiment the correlation time and the duration of the experiment are of the same order. This is the reason why FRS can address much slower particle motion than DLS. Compared to video microscopy, the length scale and time scale of the FRS experiment are about the same. However, since the holographic lattice consists of very many probe particles, the signal statistics in case of FRS are much better. For these reasons, considering the limited duration of any experiment (due to stability of the sample, the experimental setup, etc.), FRS is capable of monitoring much slower probe diffusion than all other techniques presented in this article. The shortest time accessible by FRS is determined by the duration of the bleaching laser pulse creating the holographic grating and is typically the order of 10^{-2} seconds. However, besides its strong merits, like any scattering technique FRS has one major disadvantage compared to the direct observation of the particle motion in real space using video microscopy: In the case of complex dynamics of the tracers, for example a combination of diffusive and nondiffusive modes found for tiny probes in complex heterogeneous viscoelastic media, resulting FRS signals are very difficult to interpret. Here, we should mention that the interpretation of DLS measurements in case of complex transport phenomena also may become nontrivial, as has been demonstrated above by the example of convective motion [92, 96].

Let us illustrate the potential of FRS by one recent example: frozen copolymer micelles are nanoscopic model particles consisting of a solid spherical core surrounded by a corona of dangling polymer chains. Embedded in a thermoplastic melt of polymer chains, these particles as additives largely influence the dynamical mechanical properties of the system. To investigate this effect in more detail, microrheological measurements by FRS have been employed on a system consisting of polystyrene (PS)—polyisoprene (PI) micelles (PS core, PI corona) embedded in a homopolyisoprene melt [113–117]. In this system, the micelles serve both as implicit part of the sample and as optical tracers, as already seen in some of the other experimental examples sketched above. To be suitable for the FRS experiment, the micelles have selectively been labelled within their core with a photoreactive dye. The topological architecture of this complex nanoscopic optical tracer system is sketched in Figure 6.

By measuring the self-diffusion coefficients of the micelles in various matrix polymers, which differed in their molecular weight, in dependence of micelle concentration, the characteristics of the interaction potential of the micelles (effective particle size, particle softness) could be determined. Surprisingly, it has been found that the Stokes–Einstein relation (Eq. (5)) is valid even for such a complex system: Using the zero-shear viscosity of the matrix polymer and the self-diffusion coefficients of the micelles interpolated to infinitely dilute samples, a hydrodynamic particle radius has been obtained, which fitted nicely to the topological size of the micelles as determined by X-ray and neutron scattering.

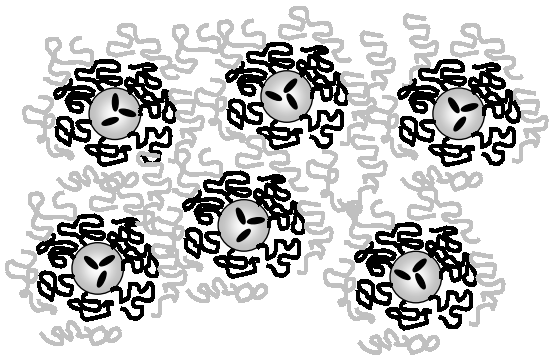


Figure 6. Copolymer micelles containing photoreactive labels (sketched as black ellipsoids within the [gray] micellar core) as tracers for microrheological studies by FRS, embedded in a homopolymer melt.

2.2.4. Optical Techniques for Microrheological Measurements in Direct Comparison

Let us close this overview on microrheological techniques using optical nanotracers by a tabellaric summary and a direct comparison (see Table 1).

The optical properties of the employed tracers strongly depend on the technique to be used, and range, as we have seen, from fluorescent or photoreactive to simply scattering light. Suitable nanoscopic tracers usually are prepared by modification of colloidal particles with organic or metallic pigments. All techniques are able to measure the quantity necessary to determine the dynamical properties of the system, and directly related to the viscoelastic moduli of the sample, the mean squared displacement $\langle \Delta r^2(t) \rangle$. The time scale of video microscopy and FRS, in case of the former, at present only limited by technically available video equipment, is insufficient to explore the dynamic mechanical spectrum $G^*(\omega)$ of a given sample in detail, especially in the regime of very short time scales (high frequencies ω). On the other hand, DWS is limited to much shorter length scales than the other methods, but is the only technique to study turbid samples.

One of the more important optical tracer techniques not discussed so far is laser deflection particle tracking (LDPT) [118], which determines the trajectory of a single diffusing colloidal probe sphere of radius several hundred nm by detecting the laser light scattered from a tightly focused beam in forward direction. If the tracer moves even slightly away from the beam's axis, the scattered light will contain an off-axis intensity asymmetry reflecting the sphere's position. One advantage of the LDPT technique is that it affords only a very small sample volume (about 5 μL). The length-scale resolution of LDPT is about 0.1 nm, at a smallest sampling time of about 10^{-4} seconds. Therefore, LDPT seems to be a powerful tool for optical microrheology. In contrast to DLS, DWS, and FRS, which only measure the mean squared displacement $\langle \Delta r^2(t) \rangle$ since they are averaging over very many probe particles at once, LDPT, like video microscopy, tracks individual particles and therefore is suited to determine the statistical distribution function of the squared particle displacements ($P[\Delta r^2(t)]$). $P[\Delta r^2(t)]$ provides more detailed information on the dynamical properties of complex samples than simply $\langle \Delta r^2(t) \rangle$, especially in the case of heterogeneous systems. Here, time and length scales of LDPT are much superior compared to current video microscopy. On the other hand, as discussed above, video microscopy is capable of tracking many particles simultaneously. This not only enables an ensemble average and thereby sharply shortens the duration of the experiment, but also allows the determination of the correlated motion of several particles as used in two-particle microrheology.

Last but not least, we would like to mention another technique based on nanotracers that are both photoreactive and fluorescent: fluorescence recovery after photobleaching (FRAP) [119–124]. This method has successfully been applied to diffusion studies in highly concentrated colloidal dispersions as well as in biological systems, and in principle works similarly to forced Rayleigh scattering (FRS). First, a periodic grating is generated by selective photobleaching of the nanoscopic tracers within the sample. Second, the time development of the amplitude of this grating is monitored not by Bragg diffraction (FRS) but by directly detecting the spatial fluorescence intensity modulation. For a more detailed description, the interested reader is referred to the literature.

In conclusion, a rich variety of optical methods employing nanoscopic tracers for microrheological measurements has been developed during the last decade. Each has its advantages and disadvantages: Some are (still) very limited in time and length scales and therefore not yet well-suited to explore the dynamic-mechanical spectrum of a viscoelastic sample. Others are restricted by the optical and structural qualities of the sample to be investigated, or require optical nanotracer particles with a special complicated architecture, which are difficult to prepare. In the near future, new technical developments and experimental strategies will further enhance our possibilities for such measurements. The necessity of optical microrheology to explore the dynamics of complex viscoelastic systems in detail at time scales not accessible by conventional dynamic-mechanical measurements exists without doubt; the careful choice of the suitable method depends on the interest of the researcher and on the sample under study.

3. OPTICAL NANOTRACERS IN DRUG DEVELOPMENT

3.1. General Remarks on Drug Development

Drug discovery is a very time consuming and expensive process—until a new product reaches the pharmacist's shelves more than 10 years have passed and over $300 million have been spent. The steadily increasing knowledge in life sciences and the rapid progress in technology and computing also influence drug discovery with new methods that promise higher efficiency. The development of a new drug is a complex process involving many different disciplines. Before the actual drug discovery, scientists focus on particular physiological processes and their associated diseases to identify biomolecular interactions suchas an enzyme

Table 1. Direct comparison of optical techniques used for microrheological measurements.

Method	Tracers	Sample	Time scale	Length scale
Video microscopy	fluorescent	transparent	10^{-3}–10^{4} s	100–10,000 nm
DLS	scattering	transparent	10^{-8}–10^{0} s	100–10,000 nm
DWS	scattering	turbid	10^{-8}–10^{0} s	0.1–10 nm
FRS	photoreactive	transparent	10^{-1}–10^{4} s	100–10,000 nm

and substrate or a receptor and ligand as therapeutic targets. With purified biomolecules, an assay test system is developed, potentially involving nanoparticles with special optical properties, allowing one to monitor the desired interaction itself as well as the influence of an active substance. This assay is used to screen a large amount of unknown artificial or natural substances, the compound library, which is an important capital of pharmaceutical companies having grown over years to a number of several hundred thousand or million compounds. After primary screening, the potentially active substances, called hits, are characterized in more complex and cellular assays and are also used in animal models. This secondary screening phase may involve several loops of chemical modifications to optimize so-called lead compounds in their physiological impact. Only five in 5,000 hits reach a stage called "Investigational New Drug" (IND), being promising enough to be registered for testing in humans where three clinical phases have to be passed. In phases I and II, the effect of the compound on a small number of healthy people and on patients is tested. Phase III involves a large number of patients for comprehensive testing on adverse effects and the best way of administering and using the drug. Only one in five compounds is finally approved: Obviously drugs being withdrawn in clinical phases constitute a substantial financial loss. In the whole process, the early steps and especially the primary screening are most easily accessible to new advances and are ideal for optimization since the testing of hundreds of thousands of compounds multiplies both effort and costs quickly [125–128].

For wet random screening, conventional techniques use large sample volumes of up to 100 μL in a carrier array of 8×12 reaction wells (96 well microtiter plate) for each test. The signal is often based on radioactive or, preferably, less hazardous fluorescent probes utilizing for example mass differences for polarization [129] or a change in total signal intensity upon reaction, and many assays involve washing steps to minimize the unspecific background. Over recent years conventional technologies have been adapted to condensed carrier formats with 384 (16×24) or even 1,536 (32×48) reaction wells of 10 μL each. However, in a small sample volume a bulk signal is often too weak [130]. Several new readout technologies have a high sensitivity, for example in assays based on energy transfer with lanthanide fluorophores, which have a long life time in the microsecond range, providing a low background in a time resolved measurement [131, 132]. Also the use of micro particles is common in drug discovery. Some well established methods use micro particles as a solid support to facilitate washing and to accumulate a signal [133]. In another sensitive assay, special beads are used that emit photons after binding to a radioactively labeled ligand [134]. Recently, nanoparticles have been used for their optical properties of scattering light, either when linked to surfaces [135] or in solution [136]. These techniques involve a special optical setup and have so far been applied to diagnostic purposes. In diagnostics, similar assay systems are used, as in drug discovery, but they are not optimized for throughput or linked into such complex logistical processes, and are therefore more flexible in sample handling and general application of new technologies.

One of the most sensitive techniques used in drug discovery is the confocal detection of single molecules. Due to a small sample volume in the focused laser beam, using piezo-dispensing the sample volume can be pushed down to 1 μL where unspecific adhesion and evaporation can still be controlled. Molecular properties of the fluorescently labeled probe are exploited using fluorescence correlation spectroscopy (FCS) [22, 137–139] or a new method called fluorescence intensity distribution analysis (FIDA) [26, 140–144]. With FIDA, the specific brightness of fluorescent molecules can be determined. Even in mixtures, populations with different specific brightness can be separated and characterized. Based on FIDA, more algorithms have been developed to use other fluorescent molecular properties suchas rotational and translational diffusion or life time. The advantage of such a highly developed technique in screening is the high information content on molecule number and characteristics, as well as on the background, so that artifacts can be identified and eliminated, improving the data quality. A typical way to set up an assay would be to determine one specific brightness of the probe bound to a target protein and another of the probe not bound, and to quantify the number of molecules in one and the other species. Thus the influence of an active compound can be measured by its ability to interfere with the probe binding to the target molecule instead. If the brightness of the fluorescent molecule is not changed upon binding and if other techniques also provide no readout, due to both biomolecules having a similar mass and being too large for polarization or FCS, nanoparticles can be applied [145–147]. After a modification suchas protein coating, this small solid support behaves like a macromolecule that accumulates fluorescently labeled molecules, creating a species with an enhanced brightness, and can thus be separated from unbound ones. The details of such an assay setup will be elucidated in the following sections.

3.2. Fluorescence Nanoparticle Assays

3.2.1. Fluorescence Intensity Distribution Analysis

For fluorescence fluctuation measurement on single molecules, a confocal microscope is used, which is equipped with lasers in the visible nanometer range and a sensitive avalanche photodiode (APD). The optics provide a pinhole and a highly focusing objective to create a confocal spot of about 1, 1, and 4 μm in x, y, and z directions. The stream of photon counts generated with such a system can be analyzed in various ways now that APDs and computational power together with underlying mathematical models are available, the most commonly known being FCS.

For FCS the fluorescence fluctuations are analyzed by calculating their autocorrelation function, which characterizes the average time the molecules are present in the confocal observation volume. Assuming that the size of the observation volume is known, it may be used together with the residence time to calculate the diffusion coefficient of the species under investigation. Still, the use of FCS in drug discovery is constrained because, upon interaction of two molecules, the mass or volume of the observed product has

to increase at least by a factor of eight to cause a twofold change in the diffusion coefficient (see Eq. (5)). Also, a change in fluorescence brightness can usually not be corrected, leading to a misinterpretation of such a reaction. FCS in combination with nanoparticles is a useful technique for diagnostics because it gives a good signal due to the large mass of the particles compared to proteins; however, for drug discovery, acquisition times are too long due to their slow diffusion. Here, FIDA provides a useful alternative.

In a typical FIDA measurement, the acquisition time is divided into time windows of equal width (usually 40 μs), and a frequency histogram (see Fig. 7) is calculated from the number of photon counts detected in each time window.

This histogram is fitted directly to give information on specific brightness and the average number of a given species within the confocal volume, involving several mathematical models and assumptions. FIDA and its related theories are described in detail elsewhere [141, 142, 144] and central points are only commented here briefly. The first point to be considered in any fluorescence intensity analysis is the spatially inhomogeneous optical excitation and detection profile, which leads to a distribution of fluorescence intensities of the fluorophore within the observation volume and therefore has to be taken into account. The excitation profile can be addressed accurately with a point spread function, but a narrowly focused laser beam is also well-described by a Gaussian–Lorentzian profile, which is multiplied with the detection profile (defined by the pinhole) to calculate the spatial brightness function. By selecting an appropriate model, adjustment parameters can be introduced that describe the spatial brightness (and volume elements with different discrete fluorescence intensities can be calculated), which can be verified by a control measurement. The main reason that fitting of the frequency histogram has not been considered a practical advance for a long time is the theory that freely floating particles located at different coordinates contribute independently to the fluorescence signal, thus requiring a direct calculation of convolutions, a mathematically clumsy and time-consuming process. A much more efficient approach is offered by the theory of generating functions, which is a convenient mathematical representation widely used in addressing problems in mathematical statistics. It has the drawback of lacking a simply understandable physical meaning, but conversely has the advantage of enabling the formulation of theories that would otherwise be too complex.

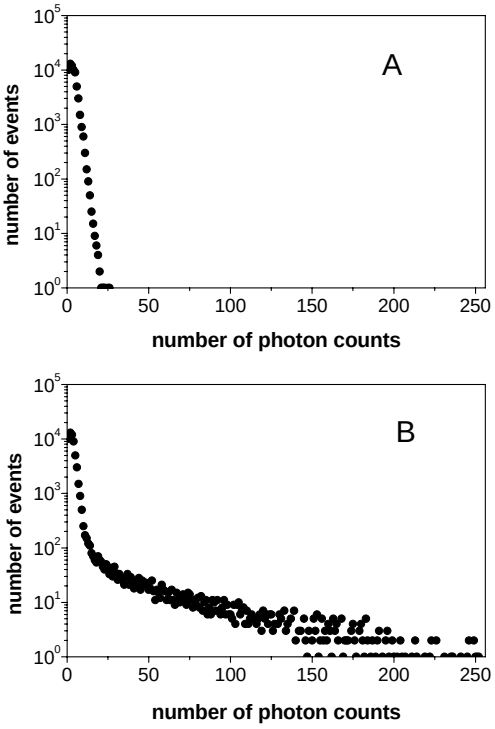


Figure 7. Photon frequency histogram of nonbound fluorescently labeled molecules (A) and of labeled molecules bound to nanoscale particles (B). The histogram shows brighter particles as events with a higher number of photon counts. A FIDA fit yields a specific molecular brightness for nonbound molecules of 7.6 ± 0.3 kHz and for the bright complexes (several labeled molecules bound to a single nanoparticle) of 1620 ± 27 kHz, indicating a loading of over 200 labeled molecules per nanoparticle.

The FIDA theory has formed the basis for the development of a series of related methods, namely two-dimensional FIDA [143] and fluorescence intensity and lifetime distribution analysis combining FIDA and fluorescence life time [142], which yield extended molecular information in addition to the molecular brightness and provide high statistical accuracy. Another intriguing but obvious approach is to divide one photon stream into different time windows to extract the specific molecular brightness and also the diffusion time, which is realized in fluorescence intensity multiple distribution analysis [144], being in a way the combination of FCS and FIDA. This new method requires longer measurement times (>2 sec) for a well-defined data set and also more extensive evaluation, therefore being mainly used for diagnostic applications also in combination with nanoparticles.

3.2.2. The Biochemistry of Nanoparticle Assays

Nanoparticles with a diameter between 50 and 200 nm used in biochemical assays must be hydrophilic and accessible to protein coating and consist usually of materials like polysaccharides or polystyrene with additional surface charges in case of the latter. To investigate nanoparticles, which have a large diffusion time and therefore would otherwise require a long measurement time, the laser beam is scanned through the sample in order to capture a sufficient number of events. There are different ways to design an assay using nanoparticles, the straightforward approach being the coupling of one binding partner to the particle and attaching a fluorescent label to the other binding partner (see Figure 8).

The sensitivity of the assay and therefore its usefulness for screening is determined by the readout sensitivity so that reagent concentrations have to be adjusted. For a FIDA fit, the average number in the confocal volume of any species should be between 0.005 and 50, which is roughly equivalent to 5 pM and 50 nM, respectively. Also, the nanoparticle should be bright enough compared to the background of free ligand so that an affinity in the low nanomolar range (10–100 fluorescent labels per nanoparticle) is required

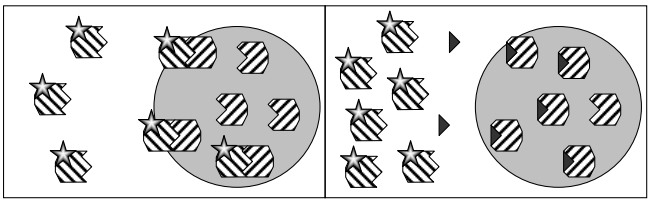


Figure 8. Nanoparticle assay where one binding partner ( ) is covalently bound to the particle, and the other binding partner ( ) is labelled with a fluorescent dye thus generating a FIDA signal when accumulating at the surface of the nanoparticle. An active compound (►) would compete for the binding sites of the target molecule located on the surface of the nanoparticle, thereby reducing the fluorescence brightness of the particle.

for the binding partners. With these constraints, a typical nanoparticle assay can be described as follows: The concentration of the nanoparticle is usually adjusted between 10 and 50 pM, avoiding any unspecific binding or light-scattering effects. Depending on the size of the coated binding partner, a number between 50 and 500 molecules can be attached to particles with a diameter of 50 nm. In an example, considering a rather large protein with a mass of 100,000 Dalton (100 kDa), it is reasonable to expect a final concentration of 500 pM of protein coated to a total of 10 pM nanoparticles (50 binding sites [ ] per nanoparticle). If the labeled binding partner ( ) has a dissociation constant (K_D) of 1 nM and is also added in this concentration, half of all coated proteins, that is 25 out of 50, are expected to be occupied with the labeled binding partners. Using in this example a specific molecular brightness of the free labeled protein of 30 kHz, the FIDA readout would give the following results (In a real experiment, results would obviously be influenced by experimental accuracy!): Species 1 corresponding to the free labeled binding partner with a specific molecular brightness of 30 kHz has a number density of 0.75 (1 nM added, roughly equivalent to an average of 1 molecule in the confocal volume; 25 labeled proteins bound per particle, 0.01 nM particles present, i.e., 0.25 nM bound, leaving 0.75 nM of free ligands). The bound species 2 would have a molecular brightness of 750 kHz (assuming 25 of 50 proteins per nanoparticle complexed with the labeled binding partner, each having a molecular brightness of 30 kHz; $25 \cdot 30$ kHz $= 750$ kHz) and a number density of 0.01 (The bright nanoparticles are measured as the second component.) To condense all parameters into one meaningful number, the degree of binding is calculated by dividing the intensity of the bound protein, which is obtained by multiplying the species number density with its specific molecular brightness, by the total intensity of both bound and free labeled binding partner. In the example, the degree of binding would be 25% $(750 \times 0.01/(750 \times 0.01 + 30 \times 0.75) = 0.25)$. For screening measurements, the analysis would be stabilized and the number of free parameters would be reduced by determining the molecular brightness of both species in control measurements and fixing it for subsequent sample measurements. The decrease of the binding degree can then quickly be determined in samples with compounds binding to the free labeled protein and thus preventing it from accumulating to the nanoparticle.

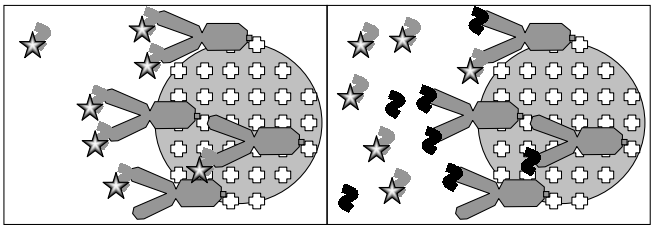


Figure 9. Competition assay where the labeled probe ( ) is complexed to the nano particle via a biotinylated (■) fix antibody ( ) linked to coated Streptavidin ( ). The analyte ( ) will compete for the binding site, which is detected as a lower specific particle brightness in FIDA measurements.

For drug discovery or diagnostics, it is always favorable to establish generic assay systems to minimize effort and risk. This can be done with nanoparticle assays by using a standard protein for coating so that the particles do not have to be modified for each individual assay. An appropriate coating protein also used for several conventional applications is streptavidin, which is found in egg white and binds biotin, an essential vitamin, with exceptionally high affinity most likely to prevent microbial growth. By attaching one or more biotin molecules to a protein, it can easily be linked to the nanoparticles. Other tools commonly used for a more advanced assay design are monoclonal antibodies: For a number of proteins that play an important physiological role, several antibodies directed to different distinct parts of the protein are available. To link an antibody, either streptavidin-coated nanoparticles and biotinylation can be used or the tight specific binding of a special antibody part to Protein A. Antibodies are ideally used to set up analytical assays, following a competition or a sandwich binding principle. In a competition assay (see Figure 9), the protein recognized by the antibody is labeled to function as a probe and is used to generate the FIDA signal. If the analyte, that is, the protein in its natural unlabeled form, is present the binding degree of the probe will decrease in a concentration-dependent manner.

In the sandwich assay (see Figure 10), another binding step is introduced, and the probe is replaced by a second monoclonal antibody, which is labeled. The analyte can now bind to the first antibody linked to the nanoparticle and is itself bound by the second labeled antibody, thus accumulating its fluorescence to the nanoparticle and generating a higher binding degree.

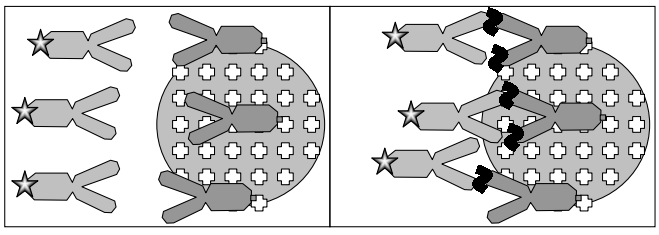


Figure 10. Sandwich assay where the nanoparticle has a high brightness if the analyte ( ) is present, so that the second antibody carrying a fluorescent label can be accumulated.

In general, a so-called positive or increasing signal, as is the case for the sandwich assay, is preferred, being less prone to artifacts. On the other hand, in this case each additional binding partner is required to have a high affinity in order to generate a high binding degree.

4. CONCLUDING REMARKS

We have tried to review the use of nanoscopic optical tracers both in microrheology and modern drug discovery. It has been shown that nanoparticles with special optical properties, suchas fluorescence, photoreactivity, or strong light-scattering power, are very useful to access detailed information in the field of soft condensed matter and biological systems. Here, the possibilities of the new method of microrheology, that is, determination of viscoelastic properties of complex fluids by tracking of small tracer particles, will develop even further in the future with technical improvements such as high-speed CCD cameras, high-speed image recording, etc. On the other hand, fluorescence techniques in combination with fluorescent nanoparticles will continue to play an important role in modern drug discovery, since they provide rapid access to significant signals from a very small number of probe particles.

GLOSSARY

Assay A controlled experiment for the quantitative estimation of a substance by measuring its effect in a living organism (in vivo) or its reactivity with an isolated biologically active substrate (in vitro). Optical nanotracers can be employed to enhance the experimental signal for such quantification using special fluorescence techniques.

Brownian motion Random motion of nanoscopic particles embedded in a fluid medium, driven by thermally incited density fluctuations of the medium. The particle trajectories are characterized by a "random walk" pattern, with the particle mean squared displacement in dependence of time t given as $\langle \Delta r^2(t) \rangle = 6\,Dt$ and D the self-diffusion coefficient.

Generalized Stokes-Einstein equation $\widetilde{G}(s) = s\tilde{\eta}(s) = k_B T/\pi\, as\langle \Delta \tilde{r}^2(s) \rangle$ relates the single particle mobility, expressed by the mean squared displacement of a tracer particle $\langle \Delta \tilde{r}^2(s) \rangle$, to the viscoelastic modulus of a complex fluid $\widetilde{G}(s)$. s is the Laplace transform variable of the time t. The generalized Stokes-Einstein equation provides the basis for optical microrheology.

Multiparticle video tracking One of the most powerful techniques used in optical microrheology, since it allows one to measure the single particle trajectories of several tracer particles simultaneously. A videocamera and a videorecorder are used to record the particle positions of the tracer particles observed under an optical microscope as a function of time. Later, the picture sequence is analyzed using digital image processing.

Optical microrheology By determining the trajectories of a diffusing tracer particle embedded in a viscoelastic fluid, using optical techniques, the viscoelastic moduli of the fluid are calculated using the generalized Stokes-Einstein equation.

Viscoelasticity Behavior of complex fluids, which shows a combination of simple liquid behavior (viscous flow) and solid behavior (elasticity) usually explored using oscillatory shear experiments (mechanical spectroscopy). The viscoelasticity is described by the viscoelastic modulus $G^* = G' + iG''$, a complex number containing an elastic contribution G' and a viscous contribution G''. As described in this chapter, microrheology provides a powerful alternative to explore the viscoelastic properties of complex fluids.

REFERENCES

1. A. Kasper, E. Bartsch, and H. Sillescu, *Langmuir* 14, 5004 (1998).
2. J. Bongers, H. Manteufel, H. Versmold, and K. Vondermassen, *J. Chem. Phys.* 108, 9937 (1998).
3. A. H. Marcus and S. A. Rice, *Phys. Rev. E* 55, 637 (1997).
4. C. A. Murray and D. G. Grier, *Ann. Rev. Phys. Chem.* 47, 421 (1996).
5. W. Schaertl and H. Sillescu, *J. Colloid Interf. Sci.* 155, 313 (1993).
6. K. Bacia, I. V. Majoul, and P. Schwille, *Biophys. J.* 83, 1184 (2002).
7. Z. Foldes-Papp, U. Demel, W. Domej, and G. P. Tilz, *Exp. Biol. and Medi.* 227, 291 (2002).
8. G. Chirico, C. Fumagalli, and G. Baldini, *J. Phys. Chem. B* 106, 2508 (2002).
9. M. Bohmer, M. Wahl, H. J. Rahn, R. Erdmann, and J. Enderlein, *Chem. Phys. Lett.* 353, 439 (2002).
10. S. T. Hess, S. H. Huang, A. A. Heikal, and W. W. Webb, *Biochemistry* 41, 697 (2002).
11. J. J. Zhao, S. C. Bae, F. Xie, and S. Granick, *Macromolecules* 34, 3123 (2001).
12. A. S. Kyriakidis, S. G. Yiantsios, and A. J. Karabelas, *J. Colloid Interf. Sci.* 195, 299 (1997).
13. R. Klein and G. Nagele, *Curr. Opin. in Colloid In.* 1, 4 (1996).
14. D. Asnaghi, M. Carpineti, M. Giglio, and A. Vailati, *Physica A* 213, 148 (1995).
15. E. Bartsch, V. Frenz, F. Renth, and H. Sillescu, *Macromol. Symp.* 79, 109 (1994).
16. P. N. Pusey, A. D. Pirie, and W. C. K. Poon, *Physica A* 201, 322 (1993).
17. K. Schatzel, *Adv. Colloid Interf. Sci.* 46, 309 (1993).
18. R. Piazza and V. Degiorgio, *J. Phys.: Condens. Matter* 5, B173 (1993).
19. W. Hartl, H. Versmold, U. Wittig, and P. Linse, *J. Chem. Phys.* 97, 7797 (1992).
20. W. Hartl and H. Versmold, *Zeitschrift Fur Physikalische Chemie–Wiesbaden* 139, 247 (1984).
21. T. Basche, *Single Mol.* 2, 237 (2001).
22. Z. H. Zhong, A. Pramanik, K. Ekberg, O. T. Jansson, H. Jornvall, J. Wahren, and R. Rigler, *Diabetologia* 44, 1184 (2001).
23. T. Basche and A. Mews, *Abstracts of Papers of the American Chemical Society* 221, 52-PHYS (2001).
24. M. Rudiger, U. Haupts, K. J. Moore, and A. J. Pope, *J. Biomol. Screen.* 6, 29 (2001).
25. L. Edman and R. Rigler, *Proc. Natl. Acad. Sci. U.S.A.* 97, 8266 (2000).
26. P. Kask, K. Palo, D. Ullmann, and K. Gall, *Proc. Natl. Acad. Sci. U.S.A.* 96, 13756 (1999).
27. T. Winkler, U. Kettling, A. Koltermann, and M. Eigen, *Proc. Natl. Acad. Sci. U.S.A.* 96, 1375 (1999).
28. T. G. Mason and D. A. Weitz, *Phys. Rev. Lett.* 74, 1250 (1995).
29. T. G. Mason, T. Gisler, K. Kroy, E. Frey, and D. A. Weitz, *J. Rheol.* 44, 1433 (2000).
30. T. G. Mason, T. Gisler, K. Kroy, E. Frey, and D. A. Weitz, *J. Rheol.* 44, 917 (2000).
31. T. G. Mason, H. Gang, and D. A. Weitz, *J. Opt. Soc. Am. A* 14, 139 (1997).

32. T. G. Mason, H. Gang, and D. A. Weitz, *J. Mol. Struct.* 383, 81 (1996).
33. T. G. Mason and D. A. Weitz, *Phys. Rev. Lett.* 75, 2770 (1995).
34. T. G. Mason, J. Bibette, and D. A. Weitz, *Phys. Rev. Lett.* 75, 2051 (1995).
35. M. G. McPhie, P. J. Daivis, I. K. Snook, J. Ennis, and D. J. Evans, *Physica A* 299, 412 (2001).
36. T. Srokowski, *Phys. Rev. E* 6403, no. 031102 (2001).
37. B. G. de Grooth, *Am. J. Phys.* 67, 1248 (1999).
38. P. Schram and I. P. Yakimenko, *Physica A* 260, 73 (1998).
39. M. Hernandez-Contreras, M. Medina-Noyola, and A. Vizcarra-Rendon, *Physica A* 234, 271 (1996).
40. K. G. Wang, *J. Phys. A: Math. Gen.* 27, 3655 (1994).
41. K. G. Wang, L. K. Dong, X. F. Wu, F. W. Zhu, and T. Ko, *Physica A* 203, 53 (1994).
42. J. D. Ferry, *Rev. Mod. Phys.* 31, 130 (1959).
43. M. J. Solomon and Q. Lu, *Curr. Opin. Colloid In.* 6, 430 (2001).
44. G. H. Koenderink and A. P. Philipse, *Langmuir* 16, 5631 (2000).
45. A. J. Banchio, G. Nagele, and J. Bergenholtz, *J. Chem. Phys.* 111, 8721 (1999).
46. A. L. Kholodenko and J. F. Douglas, *Phys. Rev. E* 51, 1081 (1995).
47. G. D. J. Phillies, *Macromolecules* 17, 2050 (1984).
48. A. J. Levine and T. C. Lubensky, *Phys. Rev. E* 6501, art. no. 011501 (2002).
49. J. C. Crocker, M. T. Valentine, E. R. Weeks, T. Gisler, P. D. Kaplan, A. G. Yodh, and D. A. Weitz, *Phys. Rev. Lett.* 85, 888 (2000).
50. A. Mukhopadhyay and S. Granick, *Curr. Opin. Colloid Int.* 6, 423 (2001).
51. Y. Song, M. Srinivasarao, and A. Tonelli, *Macromol. Symp.* 148, 395 (1999).
52. J. T. Fredrich, *Phy. Chem. Earth. A* 24, 551 (1999).
53. A. Ljunglof and J. Thommes, *J. Chromatogr., A* 813, 387 (1998).
54. T. Kraft, M. Messerli, B. Rutishauser, T. Wallimann, J. C. Perriard, and B. Brenner, *J. Muscle Res. Cell Mot.* 15, 199 (1994).
55. T. Kraft, M. Messerli, B. Rutishauser, J. C. Perriard, T. Wallimann, and B. Brenner, *Biophys. J.* 64, A26 (1993).
56. F. Koberling, A. Mews, G. Philipp, U. Kolb, I. Potapova, M. Burghard, and T. Basche, *Appl. Phys. Lett.* 81, 1116 (2002).
57. D. Gerion, F. Pinaud, S. C. Williams, W. J. Parak, D. Zanchet, S. Weiss, and A. P. Alivisatos, *J. Phys. Chem.* B 105, 8861 (2001).
58. W. Chen, Z. J. Lin, Z. G. Wang, Y. Xu, and L. Y. Lin, *J. Mater. Sci. Technol.* 14, 389 (1998).
59. V. S. Gurin and M. V. Artemyev, *J. Crys. Growth* 138, 993 (1994).
60. W. Chen, Y. Xu, Z. J. Lin, Z. G. Wang, and L. Y. Lin, *Solid State Commun.* 105, 129 (1998).
61. H. Weller, *Philos. Trans. R. Soc. London, Ser. A* 354, 757 (1996).
62. M. R. Young and P. D. Hart, *Exp. Cell Res.* 164, 199 (1986).
63. J. A. Davidson and H. S. Haller, *J. Colloid Inter. Sci.* 55, 170 (1976).
64. A. K. Dutta, K. Kamada, and K. Ohta, *J. Photochem. Photobio. A* 93, 57 (1996).
65. W. J. Brown, T. R. Sullivan, and P. Greenspan, *Histochemistry* 97, 349 (1992).
66. P. Greenspan, E. P. Mayer, and S. D. Fowler, *J. Cell Biol.* 99, A58 (1984).
67. W. Schaertl and H. Sillescu, *J. Stat. Phys.* 74, 687 (1994).
68. G. Szamel and J. A. Leegwater, *Phys. Rev. A* 46, 5012 (1992).
69. A. Vanblaaderen, J. Peetermans, G. Maret, and J. K. G. Dhont, *J. Chem. Phys.* 96, 4591 (1992).
70. B. Cichocki and B. U. Felderhof, *Phys. Rev. A* 44, 6551 (1991).
71. N. Yoshida, *Chem. Phys. Lett.* 101, 555 (1983).
72. W. Schartl, *Adv. Mater.* 12, 1899 (2000).
73. C. Graf, W. Schartl, K. Fischer, N. Hugenberg, and M. Schmidt, *Langmuir* 15, 6170 (1999).
74. D. F. Nicoli, D. C. McKenzie, and J. S. Wu, *Am. Lab.* 23, 32 (1991).
75. H. G. Barth and S. T. Sun, *Anal. Chem.* 63, R1 (1991).
76. J. C. Thomas, *J. Colloid Interf. Sci.* 117, 187 (1987).
77. A. Eshuis, G. Harbers, D. J. Doornink, and P. F. Mijnlieff, *Langmuir* 1, 289 (1985).
78. V. Degiorgio and R. Piazza, *Curr. Opin. Coll. In.* 1, 11 (1996).
79. E. Killmann, P. Sapuntzjis, and H. Maier, *Makromol. Chem., Macromol. Symp.* 61, 42 (1992).
80. G. Maret, *Curr. Opin. Coll. In.* 2, 251 (1997).
81. D. A. Weitz, J. X. Zhu, D. J. Durian, H. Gang, and D. J. Pine, *Phys. Scr. T.* 49B, 610 (1993).
82. D. S. Horne and C. M. Davidson, *Coll. Surf. A* 77, 1 (1993).
83. D. J. Pine, D. A. Weitz, J. X. Zhu, and E. Herbolzheimer, *Journal De Physique* 51, 2101 (1990).
84. D. A. Weitz, D. J. Pine, P. N. Pusey, and R. J. A. Tough, *Phys. Rev. Lett.* 63, 1747 (1989).
85. F. C. Mackintosh and S. John, *Phys. Rev. B* 40, 2383 (1989).
86. D. J. Pine, D. A. Weitz, P. M. Chaikin, and E. Herbolzheimer, *Phys. Rev. Lett.* 60, 1134 (1988).
87. S. R. Williams and W. van Megen, *Phys. Rev. E* 6404, 041502 (2001).
88. S. Kirsch, A. Doerk, E. Bartsch, H. Sillescu, K. Landfester, H. W. Spiess, and W. Maechtle, *Macromolecules* 32, 4508 (1999).
89. W. van Megen, T. C. Mortensen, S. R. Williams, and J. Muller, *Phys. Rev. E* 58, 6073 (1998).
90. E. Bartsch, V. Frenz, J. Baschnagel, W. Schartl, and H. Sillescu, *J. Chem. Phys.* 106, 3743 (1997).
91. E. Bartsch, V. Frenz, and H. Sillescu, *J. Non-Cryst. Solids* 172, 88 (1994).
92. W. Schaertl and C. Roos, *Phys. Rev. E* 60, 2020 (1999).
93. L. M. Liz-Marzán and A. P. Philipse, *J. Colloid Interf. Sci.* 176, 459 (1995).
94. S. G. Stanton, R. Pecora, and B. S. Hudson, *J. Chem. Phys.* 75, 5615 (1981).
95. C. Roos, M. Schmidt, J. Ebenhoch, F. Baumann, B. Deubzer, and J. Weis, *Adv. Mater.* 11, 761 (1999).
96. W. G. Schartl, C. Roos, and K. Gohr, *J. Chem. Phys.* 108, 9594 (1998).
97. H. M. Wyss, S. Romer, F. Scheffold, P. Schurtenberger, and L. J. Gauckler, *J. Colloid Interf. Sci.* 241, 89 (2001).
98. C. Urban, S. Romer, F. Scheffold, and P. Schurtenberger, *Macromol. Symp.* 162, 235 (2000).
99. B. R. Dasgupta, S. Y. Tee, J. C. Crocker, B. J. Frisken, and D. A. Weitz, *Phys. Rev. E* 65, art. no. 051505 (2002).
100. T. Narita, A. Knaebel, J. P. Munch, and S. J. Candau, *Macromolecules* 34, 8224 (2001).
101. A. Palmer, J. Y. Xu, S. C. Kuo, and D. Wirtz, *Biophys. J.* 76, 2317 (1999).
102. A. Palmer, T. G. Mason, J. Y. Xu, S. C. Kuo, and D. Wirtz, *Biophys. J.* 76, 1063 (1999).
103. J. Y. Xu, V. Viasnoff, and D. Wirtz, *Rheol. Acta* 37, 387 (1998).
104. V. Viasnoff, F. Lequeux, and D. J. Pine, *Rev. Sci. Instrum.* 73, 2336 (2002).
105. J. L. Harden and V. Viasnoff, *Curr. Opin. in Coll. In.* 6, 438 (2001).
106. F. Scheffold, S. E. Skipetrov, S. Romer, and P. Schurtenberger, *Phys. Rev. E* 6306, 061404 (2001).
107. L. Cipelletti and D. A. Weitz, *Rev. Sci. Instrum.* 70, 3214 (1999).
108. S. Kirsch, V. Frenz, W. Schartl, E. Bartsch, and H. Sillescu, *J. Chem. Phys.* 104, 1758 (1996).
109. A. P. Y. Wong and P. Wiltzius, *Rev. Sci. Instrum.* 64, 2547 (1993).
110. M. Antonietti, J. Coutandin, R. Grutter, and H. Sillescu, *Macromolecules* 17, 798 (1984).
111. H. Hervet, L. Leger, and F. Rondelez, *Phys. Rev. Lett.* 42, 1681 (1979).
112. D. W. Pohl, S. E. Schwarz, and V. Irniger, *Phys. Rev. Lett.* 31, 32 (1973).
113. K. Gohr and W. Schartl, *Macromolecules* 33, 2129 (2000).
114. K. Gohr, T. Pakula, K. Tsutsumi, and W. Schartl, *Macromolecules* 32, 7156 (1999).

115. W. Schartl, *Macromol. Chem. Phys.* 200, 481 (1999).
116. H. Watanabe, T. Sato, K. Osaki, M. W. Hamersky, B. R. Chapman, and T. P. Lodge, *Macromolecules* 31, 3740 (1998).
117. W. Schaertl, K. Tsutsumi, K. Kimishima, and T. Hashimoto, *Macromolecules* 29, 5297 (1996).
118. I. M. Peters, B. G. de Grooth, J. M. Schins, C. G. Figdor, and J. Greve, *Rev. Sci. Instrum.* 69, 2762 (1998).
119. T. E. Starr and N. L. Thompson, *Biophys. Chem.* 97, 29 (2002).
120. B. Fong, W. Stryjewski, and P. S. Russo, *J. Colloid Interf. Sci.* 239, 374 (2001).
121. M. Arrio-Dupont, G. Foucault, M. Vacher, P. F. Devaux, and S. Cribier, *Biophys. J.* 78, 901 (2000).
122. S. Kluijtmans, G. H. Koenderink, and A. P. Philipse, *Phys. Rev. E* 61, 626 (2000).
123. S. Kluijtmans and A. P. Philipse, *Langmuir* 15, 1896 (1999).
124. P. Wahl, *Biophys. Chem.* 57, 225 (1996).
125. J. Wolcke and D. Ullmann, *Drug Discov. Today* 6, 637 (2001).
126. R. P. Hertzberg and A. J. Pope, *Curr. Opin. Chem. Bio.* 4, 445 (2000).
127. J. J. Burbaum, *Drug Discov. Today* 3, 313 (1998).
128. W. Harrison, *Drug Discov. Today* 3, 343 (1998).
129. M. S. Nasir and M. E. Jolley, *Comb. Chem. High T. Screen.* 2, 177 (1999).
130. A. J. Pope, U. M. Haupts, and K. J. Moore, *Drug Discov. Today* 4, 350 (1999).
131. G. W. Mellor, M. N. Burden, M. Preaudat, Y. Joseph, S. B. Cooksley, J. H. Ellis, and M. N. Banks, *J. Biomol. Screen.* 3, 91 (1998).
132. G. Mathis, *Clin. Chem.* 41, 1391 (1995).
133. M. E. Jolley, C. H. J. Wang, S. J. Ekenberg, M. S. Zuelke, and D. M. Kelso, *J. Immunol. Methods* 67, 21 (1984).
134. M. Beveridge, Y. W. Park, J. Hermes, A. Marenghi, G. Brophy, and A. Santos, *J. Biomol. Screen.* 5, 205 (2000).
135. S. Kubitschko, J. Spinke, T. Bruckner, S. Pohl, and N. Oranth, *Anal. Biochem.* 253, 112 (1997).
136. P. Montagne, P. Varcin, M. L. Cuilliere, and J. Duheille, *Bioconjugate Chem.* 3, 187 (1992).
137. A. C. Lagerkvist, Z. Foldes-Papp, M. A. A. Persson, and R. Rigler, *Protein Sci.* 10, 1522 (2001).
138. R. Rigler, *J. Biotechnol.* 41, 177 (1995).
139. J. Widengren and R. Rigler, *Cell. Mol. Biol.* 44, 857 (1998).
140. S. Schaertl, S. Kraemer, B. Buda, E. Lopez-Calle, and P. Kask, *Biophys. J.* 78, 2267 (2000).
141. K. Gall, K. Palo, L. Brand, C. Eggeling, and P. Kask, *Biophys. J.* 82, 1728 (2002).
142. K. Palo, L. Brand, C. Eggeling, S. Jager, P. Kask, and K. Gall, *Biophys. J.* 83, 605 (2002).
143. P. Kask, K. Palo, N. Fay, L. Brand, U. Mets, D. Ullmann, J. Jungmann, J. Pschorr, and K. Gall, *Biophys. J.* 78, 1703 (2000).
144. K. Palo, U. Metz, S. Jager, P. Kask, and K. Gall, *Biophys. J.* 79, 2858 (2000).
145. T. Soukka, H. Harma, J. Paukkunen, and T. Lovgren, *Anal. Chem.* 73, 2254 (2001).
146. S. Schaertl, F. J. Meyer-Almes, E. Lopez-Calle, A. Siemers, and J. Kramer, *J. Biomol. Screen.* 5, 227 (2000).
147. S. Schultz, J. Mock, D. R. Smith, and D. A. Schultz, *J. Clin. Ligand Assay* 22, 214 (1999).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nano-spintronics for Data Storage

Yihong Wu

National University of Singapore, Singapore and Data Storage Institute, Singapore

CONTENTS

1. INTRODUCTION

Nanotechnology is one of the hottest research fields in the science and engineering arena. Of the many potential applications of nanotechnology, one of the most promising ones is in data storage, particularly the hard disk drives. This is because the physical size of the recording bits of hard disk drives is already in the nanometer regime, and continues shrinking due to the ever-increasing demand for higher recording densities. If the pace of areal density increase is maintained at the current level for about ten years, the dimension of the recording bit will reach the sub-10 nm regime. At this level, both the writing and reading processes will become extremely challenging, if not impossible.

The rapid shrinkage of bit size poses formidable challenges to the read sensors. This is because the sensor must be made smaller or at least comparable to the bit size, and at the same time, its sensitivity must be improved continuously so as to compensate the loss in signal-to-noise ratio due to the decrease in the bit size. The former has to rely heavily on the advance of nanotechnology, and the latter on an emerging field called spintronics [1–4]. The combination of these two fields has played an important role in advancing the areal density of magnetic recording from a few gigabits/in^2 to the current level of more than 100 Gbits/in^2 [5–8]. In addition to hard disk drives, the technologies developed have also been applied to magnetic random access memories (MRAMs) [9–11]. Further advance in the two fields is the key to realizing terabits/in^2 hard disk drives and gigabit nonvolatile memories within this decade. This chapter is intended to provide an updated review of nanometer-scale spintronics for data storage applications, with emphasis on the application of metal-based spintronics in magnetic sensors and memories.

To have a clear picture of the scope of this review, Figure 1 illustrates how spintronics is positioned in the hierarchy of various different types of data storage technologies. The data storage devices can be categorized into three major groups, that is, magnetic and optical data storage, and solid-state memory. Each of them can be divided further into several subgroups, and here, we only focus on magnetic data storage and solid-state memory. Among the magnetic storage devices, the hard disk drive (HDD) is the dominant secondary mass storage device for computers, and very likely also for home electronic products in the near future. The HDD is an integration of many key technologies, including head, medium, head-disk interface (HDI), servo, channel coding/decoding, and electromechanical and electromagnetic devices. Among them, the read head is the only component that has experienced the most changes, including some revolutionary ones in terms of both the operating principle and the structural design and fabrication processes during the last decade. The ever-increasing demand for higher areal densities has driven the read head evolving from a thin-film inductive head to an anisotropic magnetoresistive (AMR) head, and recently, the giant magnetoresistive spin-valve (GMR SV) head. There are two different forms of GMR SVs, depending on whether the sense current flows in the film plane (CIP) or perpendicular to the film plane (CPP). Currently, CIP SVs are dominant, but CPP SVs are more promising for future extremely high-areal-density heads. In addition to CPP spin valves, a magnetic tunnel junction (MTJ) with low junction resistance is also one of the potential candidates for future read sensors. The MTJ is particularly useful for MRAM applications.

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (493–544)

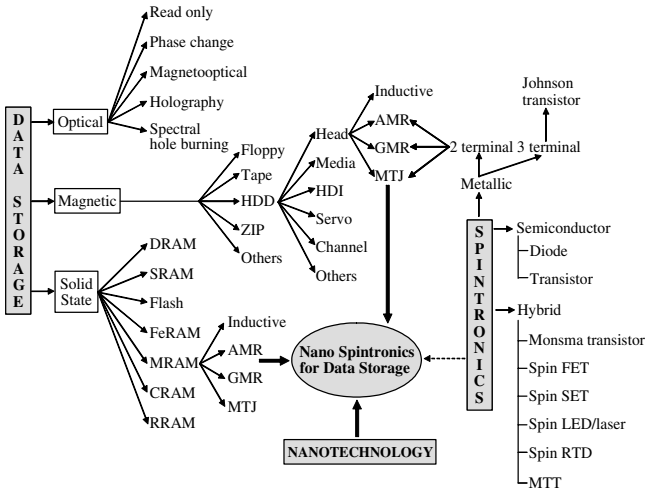


Figure 1. Spintronics and its applications in data storage and electronics. The acronyms are: DRAM—dynamic random access memory, SRAM—static random access memory, FeRAM—ferroelectric random access memory, MRAM—magnetic random access memory, CRAM—chalcogenide random access memory, RRAM—resistance random access memory, ZIP—zip drive, HDI—head–disk interface, AMR—anisotropic magnetoresistance, GMR—giant magnetoresistance, MTJ—magnetic tunnel junction, FET–field-effect transistor, SET—single-electron transistor, LED—light-emitting diode, RTD—resonant tunneling diode, MTT—magnetic tunnel transistor.

As is with the case of read sensors, the MRAM started from the ferrite core design to AMR, GMR pseudospin valve, and the current MTJ cell design. However, it is only after the introduction of MTJ that the MRAM has taken off, and is expected to be in mass production in a few years' time. The GMR and MTJ sensors are generally called magnetoelectronic or metal-based spintronic devices, which are the focuses of this review. The metal-based spintronic devices are based primarily on the spatial modulation of electron spins through using layered structures of magnetic and nonmagnetic materials. The lack of capability in charge modulation in these types of structures may eventually limit their ultimate performances in terms of both the magnetoresistance and other functionalities. To address this issue, recently, a great deal of effort has been devoted to the development of magnetic semiconductors which allow the modulation of both the spins and charges. The advance in this field may eventually lead to spintronic devices with performances superior to their metal-based counterparts. In addition to pure metal-based or semiconductor-based spintronic devices, hybrid devices making use of both technologies also have been explored actively in recent years.

The organization of the rest of this chapter is as follows. Section 2 gives a brief introduction to the basics of magnetic data storage and spintronics, and explains why both the spintronics and nanotechnology are crucial for magnetic data storage. Section 3 discusses different types of magnetoresistance effects, and their applications in read sensors. In-depth reviews of the magnetic and electronic aspects of spin valves are given in Sections 4 and 5, respectively. Section 6 deals with the MTJs, with their applications in MRAMs discussed in Section 7. Section 8 touches briefly on the semiconductor spintronics. Finally, we discuss the future trend of data storage and the role of spintronics in future data storage in Section 9.

2. BASICS OF MAGNETIC DATA STORAGE AND SPINTRONICS

2.1. Magnetic Recording and Magnetic Memory

The hard disk drive at the system level is the integration of many key technologies and components, which include the storage medium, read/write transducer, channel coding/decoding, servocontrol, head/disk interface, tribology, and electromechanical and electromagnetic systems [12–14]. In this subsection, however, we focus on the physics and materials aspects of magnetic recording, that is, topics related to the storage medium and read/write heads [15].

The thin-film medium for a hard disk drive is a stack of multiple layer thin films formed on either an NiP-coated aluminum alloy or a glass substrate [16]. Among the multiple layers of thin films, the early hard disk media only employed a single magnetic layer as the recording layer, which is usually a polycrystalline alloy of Co, Cr, and Pt with additional elements such as Ta or B to improve the magnetic properties. The latest magnetic media, however, employ more than one magnetic layer as the recording layer, such as the antiferromagnetically coupled (AFC) media to improve the thermal stability of the information bits recorded on the hard disk [17–19]. When recording the information onto the disk, a writer which can generate a sufficiently high magnetic field is used to switch the magnetization of a localized area of the media to one of two fixed directions, with one of them representing digit "1" and the other representing digit "0." Each of the localized areas can be considered as one bit. Each bit consists of a number of partially exchange-coupled magnetic grains, as shown schematically in Figure 2.

When reproducing the information from the disk, the early disk drive employs the same write head as the read head, that is, the signal is detected by measuring the voltage change across the coils induced by the flux variation when

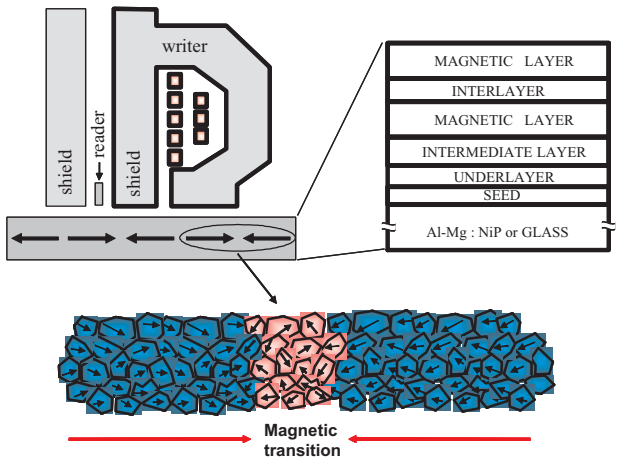


Figure 2. Schematic of a magnetic recording head and media. The head typically consists of an inductive writer and an AMR or GMR reader.

the head passes by the recorded media. Although the performance of the inductive head had been improved continuously to meet the requirement of rapidly increasing areal density, its structural design and working principle did pose certain limitations to the ultimate performance of this type of head. That is the reason why, in the beginning of the 1990s, a new type of head which is based on the anisotropy magnetoresistance of certain magnetic materials was introduced by IBM, and later also adopted by other head and drive companies. This type of head is widely called an AMR head [20]. The AMR effect stems from the spin–orbit interaction of electrons, and thus it normally appears in magnetic materials; its value, however, varies from less than 1% to a few percent, depending on the type of material. The common material that is used for an AMR head is Permalloy, of which the MR ratio is about 2–4%. It decreases almost linearly with thickness for ultrathin films (<10 nm).

Although the AMR head showed much better performance in terms of sensitivity and scalability than the thin-film inductive head, its pace of structural miniaturization and performance improvement was soon outpaced by the rapid advance in areal density. The rapid increase of areal density requires simultaneously a decrease in both the bit length and the track width. The former requires a thinner and more sensitive sensor, while the latter requires a large magnetoresistance (MR). Unfortunately, it is difficult for the AMR head to satisfy both requirements simultaneously because the MR ratio of the AMR head decreases rapidly when its thickness is reduced to less than 10 nm [21]. On the other hand, in 1988, a much larger MR effect, that is, GMR, was discovered in artificially made magnetic and nonmagnetic multiple layers [22, 23]. The MR ratio of this type of superlattice structure is more than one order of magnitude larger than that of the AMR. The GMR structure, however, could not be applied directly to the read head as it was because of the large magnetic field that is required to switch the magnetization of the magnetic layers from one direction to the other. This has prompted IBM to invent a more practical structure for a read sensor which is called a spin valve [24].

The state-of-the-art spin valve consists of a dozen thin layers, the heart of which is a trilayer structure consisting of two ferromagnetic layers separated by a nonmagnetic spacer, which is usually copper. The signal detection principle is based on the fact that the resistance of the trilayer is dependent on the relative orientation of the magnetization of the two ferromagnetic layers. The resistance is high when the magnetizations of the two layers are in opposite directions, and low when they are in the same direction. To have a linear response from the sensor, the angle between the two magnetizations is normally set at 90° at zero field, with one of them "pinned" at a direction perpendicular to the media surface through exchange coupling with an antiferromagnet, and the other free to rotate in response to the fringe field of magnetic transitions recorded on the magnetic media.

There are two different forms of spin-valve sensors, depending on whether the current flows in the plane of the stack of layers or perpendicular to them. The former is called a current-in-plane, or CIP, spin-valve sensor, and the latter a current-perpendicular-to-plane, or CPP, sensor [25]. So far, CIP is dominant, but CPP is expected to play an important role in future terabit recording systems. An alternative to the CPP spin valve is the magnetic tunnel junction, or MTJ, in which the current also flows perpendicular to the plane [26–29]. The major difference between the CPP spin valve and the MTJ is that the latter is composed of two ferromagnetic layers separated by an insulator instead of a metal. Therefore, the electrical conduction in MTJ is based on quantum-mechanical tunneling. The MTJ is attractive because its MR ratio is generally more than double that of spin valves. However, the primary drawback of MTJ is that its junction resistance is generally larger than that of the CIP or CPP sensors, which may affect its performance as a read sensor due to the increase in thermal noise and the decrease in bandwidth when the junction size is reduced to a certain value. In addition to this, the quantum-mechanical nature of electron transport also causes shot noise which increases with the intensity of the output signal. Thus, the major application of MTJ so far is in MRAM. As one of the newcomers in the memory hierarchy, the MRAM has the potential to replace some of the commercially available memories in the near future.

Except for the inductive read head, all other read sensors discussed above make use of both the charge and spin properties of electrons. AMR stems from spin–orbit interaction, while the spin valve and MTJ operate based on the spin-dependent scattering and tunneling both inside the magnetic layers and at the magnet/nonmagnet or insulator interfaces. In contrast to Si-based electronics, the GMR-based sensors are often referred to as a subcategory of an emerging field called spintronics (see Fig. 1).

2.2. Spintronics

Electrons possess both charges and spins. The motion of charges forms the current. The ability to control or modulate the charge transport has made it possible to form functional devices such as diodes and transistors. This is so far only possible in semiconductors instead of metals because the latter have too many electrons per unit volume; the variation of charge distribution, if any, is limited to a few atomic layers at the surface that can hardly cause any measurable change in the conductance of the metal. Therefore, functional electronic devices have not been realized in pure metals with dimensions larger than the mean-free path of electrons. Although they belong to the same family of metals, metallic magnetic materials have another additional degree of freedom which can be used to vary their electronic transport properties—the spin of electrons. As the spin of electrons in magnetic materials can be easily manipulated using an external magnetic field without suffering the electrostatic screening effect as the charges do when they are subjected to an electric field, it is possible to alter the conductivity of magnetic materials without changing the carrier distribution inside the material. This forms the basis of GMR-based electronics or sometimes is also called magnetoelectronics. It is a subfield of spintronics or spinelectronics [1–4].

The spintronics in its broader sense contains all types of electronics that make use of both charges and spins. In semiconductor-based electronics, we only control the

charge motion, and ignore the spins. In fact, it is difficult to control the spins in semiconductors because they do not have specific directions under normal conditions, unless the electrons are injected into from external sources with a net polarization. On the other hand, in metallic magnetic materials, it is easy to control the spins due to the strong ferromagnetic coupling among the spins, but it is difficult to control the charges. Therefore, a question naturally arises here: Can we have a kind of material in which we can control both charges and spins? The answer is "yes," and this type of material is called a magnetic semiconductor [30–32]. The magnetic semiconductor is usually made through adding magnetic impurities to host semiconductors. It is not necessary, however, that every semiconductor can be made magnetic using this approach because some of them still do not exhibit any magnetic properties even after they are doped with a substantial amount of magnetic impurities. Some of them, although being magnetic, show a very low Curie temperature [33, 34]. However, the situation changed drastically in recent years due to the intensive efforts made by researchers in this field in many research organizations. Several different types of magnetic semiconductors having a Curie temperature higher than room temperature have been found. These include, but are not limited to, GaN:Mn [35, 36], CdGeP:Mn [37], GaP:Mn [38], TiO_2:Co [39], and ZnO:Co [40]. It should be noted, however, that all of these are based on preliminary experimental results; further experiments are required to verify the results. The progress was made not only in materials themselves, but also the application of these materials in creating new functional devices such as semiconductor-based magnetic tunnel junctions [41, 42] and spin-injection devices [43, 44]. Although the current technology for a read sensor is based on metallic spintronics, semiconductor-based spintronics has the potential to provide sensors or storage elements with superior performances for next-generation data storage devices.

Although the read sensor for magnetic recording and storage cells for magnetic memory are based on spintronics, the storage of information in disk media is still based on classic physics, and it does not involve spintronics. However, as the bit size continues decreasing, one day the bit size will approach that of atoms. At this stage, a fundamental change in the information storage principle might be required. One of the possible scenarios is to store information in the reciprocal space or energy domain. In this case, the spin of electrons and nuclei instead of the magnetization of magnetic grains will be important. This is closely related to another emerging field called quantum information storage [45].

2.3. Nanotechnology and Magnetic Data Storage

Nanotechnology is indispensable for all kinds of data storage devices, be it magnetic, optical, or solid-state devices. This is because the common driving force for any type of data storage technology is the areal density, provided that the bit-cost performance also improves with the areal density. As a rule of thumb, the bit size shrinks by one order when the areal density increases by two orders. Figure 3 shows the size of a square bit as a function of the areal density. The bit size falls into the nanometer regime (1–100 nm) when the areal density reaches 70 Gbits/in^2, which was already realized about two years ago. To write information to and read information from these small bits, the write and read heads must have a size which is smaller than or at least comparable to that of the bit. In addition to the read/write head, the spacing between the disk and the head must also be kept as small as possible, and in a practical case, it is about one fifth of the bit size. Therefore, once the areal density enters the hundreds of gigabits range, the dimensions of almost all of the key components in magnetic recording have to be in the nanometer regime.

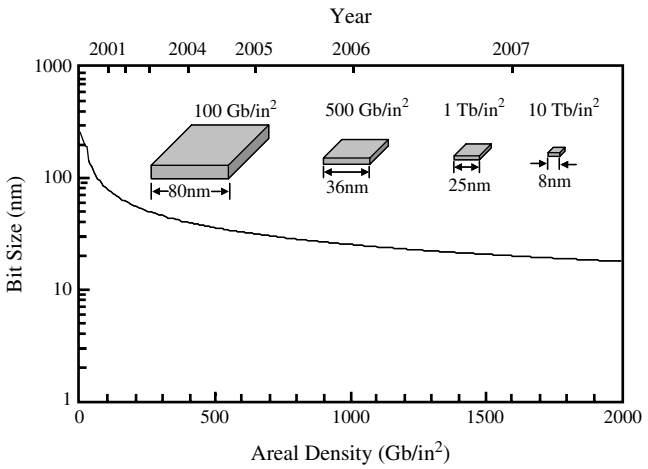


Figure 3. Bit size as a function of areal density for square bits. Also shown is the year at which the areal density will be reached by assuming a compound annual growth rate of 60%. Inset shows the size of square bits at areal densities of 100, 500, 1000, 10,000 Gbits/in^2, respectively.

The continuous shrinkage of bit size poses formidable challenges to both media and read heads. For low-areal-density recording, each single bit contains tens of thousands of small magnetic grains. At this level, one will not "feel" the difference when one or two grains are missing. However, the number of grains per bit decreases monotonically with the areal density. At 100 Gbits/in^2, the number of grains is reduced to less than 100. This results in an increase of media noise due to the particulate nature of the media. One of the possible ways to suppress the noise is to reduce the grain size of the media so that each bit will contain a greater number of grains. However, this is not so straightforward because the superparamagnetic phenomenon will appear, making the information bit thermally unstable. Although the recent advance in multiple-layered media has greatly enhanced the thermal stability of the media with small grains, it is still far from being a perfect solution. The thermal stability of the media is not only just affected by the grain size, but also very much by the grain size distribution. The latter is technically more challenging because one must use a mass-production technique which can produce not just small grains, but also monodispersed grains. This is the area where nanotechnology is expected to play an important role. One of the possible ways to realize such a kind of media is to fabricate single-domain particles or dots by using nanofabrication techniques [46–48]. Although one of the major applications of nanostructured magnetic materials and nanomagnetism is in storage media,

we will not discuss this further in this chapter because, so far, the media are still in the "passive" form, and do not involve spintronics.

Now we turn to the read sensor, which is our main focus in this chapter. We are concerned about the sensor when its size is reduced because, first, the sensor size must be reduced accordingly when the bit size decreases, and second, the output and impedance of the sensor are dependent on the sensor size. So far, the sensor fabrication has leveraged the semiconductor manufacturing processes. But this trend is going to be reversed because the size shrinkage of read sensors has outpaced the shrinkage of critical dimensions of integrated circuits [49]. In a few years' time, the dimension of the sensor will approach 50 nm, at which point there may not even be a solution for the semiconductor industry at the moment. In addition to the manufacturing issues, one also needs to find a way to improve the performance of the sensors so that the output of the sensor will not decrease significantly when its size is reduced.

The above discussion demonstrates clearly the importance of nanotechnology and spintronics in magnetic recording and memory devices. In fact, spintronics is closely related to nanotechnology because most of the phenomena related to bulk and surface magnetism have a characteristic length in the nanometer regime. In the rest of this chapter, we discuss in detail the fundamental concepts and latest development of metallic spintronics and its applications in magnetic data storage.

3. MAGNETORESISTANCE AND THE ASSOCIATED READ SENSORS

3.1. Anisotropic magnetoresistance

The resistivity of a nonmagnetic material is usually independent of the magnetic field. However, this may not be the case for magnetic materials. The phenomenon that the resistivity of a ferromagnetic material depends on the relative angle between the current and magnetization direction of the material is called the anisotropic magnetoresistance effect, or AMR, which was discovered in 1857 by Thomson [50]. A series of theoretical studies carried out a century later showed that the AMR effect stems from the spin–orbit interaction [51]. The importance of the AMR effect was recognized in the 1970s when a large AMR was found in a number of alloys based on iron, cobalt, and nickel [52]. Materials exhibiting a normal AMR effect show a maximum resistivity when the current is parallel to the magnetization direction ($\rho_{||}$) and a minimum resistivity when the current is perpendicular to the magnetization direction ($\rho_{\perp}$). A measure for the size of this effect is the AMR, ratio which is defined by

$$\mathrm{MR} = \frac{\rho_{||} - \rho_{\perp}}{\rho_{||}} \tag{1}$$

At intermediate angles between the current and magnetization direction, the resistivity of an AMR material is given by

$$\rho(\theta) = \rho_{\perp} + (\rho_{||} - \rho_{\perp}) \cos^2 \theta \tag{2}$$

where θ is the angle between the current and the magnetization direction.

3.2. AMR Sensor

The immediate application of the AMR was in magnetic recording because sensors based on the AMR effect offer higher output as compared to the thin-film inductive head [53, 54]. Although, for Ni (Fe, Co) alloys, the largest AMR effect so far was found for $Ni_{70}Co_{30}$, 26.7% at 4.2 K and 6.6% at 300 K [52], the material of choice for magnetic recording applications is permalloy because of its softness, high permeability, and low magnetorestriction. The typical AMR ratio for thin permalloy films (30–50 nm) is about 2%, although the AMR of its bulk counterpart can be as high as 4% [13, 21]. When being used as a sensor, the magnetization direction is normally set at 45° with respect to the current direction at zero field so as to maximize the sensitivity [55]. This is apparent from the fact that the first derivative of ρ is maximum when $\theta = 45°$.

Figure 4 shows a schematic drawing of an AMR sensor element. For the sake of simplicity, we assume that the external magnetic field points in the y-axis direction. Assuming that the entire element is a single-domain particle with a uniaxial anisotropy, the magnetization direction can be obtained from minimization of the energy density [13]:

$$\varepsilon = -HM_s \sin\theta + K_u \sin^2\theta + \frac{1}{2} H_d M_s \sin^2\theta \tag{3}$$

Note that here only the magnetostatic energy, anisotropy energy, and demagnetizing energy are taken into account; all other energy terms are neglected. In Eq. (3), H is the external magnetic field, M_s is the saturation magnetization, K_u is the uniaxial anisotropy energy constant, H_d is the demagnetizing field, and θ is the angle between the magnetization and easy axis direction. Minimizing the energy density gives

$$\sin\theta = \frac{H}{2K_u/M_s + H_d} = \frac{H}{H_k + H_d} \tag{4}$$

where $H_k = 2K_u/M_s$ is the anisotropy field. Substituting Eq. (4) into Eq. (2) yields

$$\rho(H) = \rho_{||} - (\rho_{||} - \rho_{\perp}) \left(\frac{H}{H_k + H_d} \right)^2 \tag{5}$$

For soft materials, $H_k \ll H_d$; therefore,

$$\rho(H) = \rho_{||} - (\rho_{||} - \rho_{\perp}) \left(\frac{H}{H_d} \right)^2 \tag{6}$$

It is apparent from the above equation that the simple AMR sensor element exhibits a nonlinear response to the external field, which cannot be used as a read sensor for magnetic recording as it is. However, it is not difficult to realize from

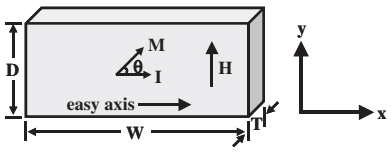


Figure 4. Schematic of a single-domain AMR sensor element. H is the external field, M is the saturation magnetization, and I is the current.

Eq. (6) that the sensor can be made linear if an additional field which is much larger than that of the external field is added to it to make the total effective external field as $H' = H + H_B$ with $H_B \gg H$. In the case of $H_d \gg H$, which is true for magnetic recording, Eq. (6) becomes

$$\rho(H) \approx \rho_{||} - (\rho_{||} - \rho_{\perp})\left(\frac{H_B}{H_d}\right)^2 - (\rho_{||} - \rho_{\perp})\frac{2HH_B}{H_d^2} \qquad (7)$$

It shows that now the sensor responds linearly to the external field. H_B is the so-called bias field or, more precisely, the traverse bias field because it is perpendicular to the easy axis direction of the sensor element. In actual sensor design, the strength of the bias field is chosen such that the magnetization direction at zero field is about 45° away from the easy axis so as to maximize the sensitivity. It is obvious from Eq. (7) that the smaller the demagnetizing field, the larger the sensitivity.

There are many different ways to form a traverse bias. Among them, the most successful was the soft adjacent layer (SAL) bias scheme, in which a soft ferromagnetic layer is laminated with the sensing layer via a thin insulating spacer, as shown schematically in Figure 5 [55–58]. As the SAL is normally chosen such that most of the current flows through the sensing layer, the magnetic field induced by the sensing current magnetizes and saturates the SAL into one direction (pointing upward in Fig. 5). The fringe field thus generated, in turn, provides a traverse bias to the sensing layer itself. The SAL scheme offers several advantages, such as adjustable bias field, relatively uniform bias field distribution, and reduced demagnetizing field. Although it also has drawbacks such as the current shunting effect, it so far has remained the most successful engineering design. In actual sensors, in addition to the traverse bias, one also needs a longitudinal bias to stabilize the domain structure so as to reduce the Baukhausen noise caused by the domain-wall motion [55, 56].

As we mentioned above, the AMR sensor, intrinsically, is not a linear sensor. In addition to the nonlinearity issue, it also suffers drawbacks such as thermal asperity [59] and side reading asymmetry [60, 61]. Perhaps the most fatal shortcoming of the AMR sensor is that it is difficult to scale it down in thickness so as to meet the requirement of shrinkage in bit length and reduction in $M_r t$ of the media [21]. These intrinsic characteristics of the AMR plus the tremendous progress made in GMR in the early 1990s have determined the short lifetime of the AMR sensor in the history of hard disk drives. It was gradually replaced by the spin-valve sensor, which was first introduced into disk drives by IBM in 1997.

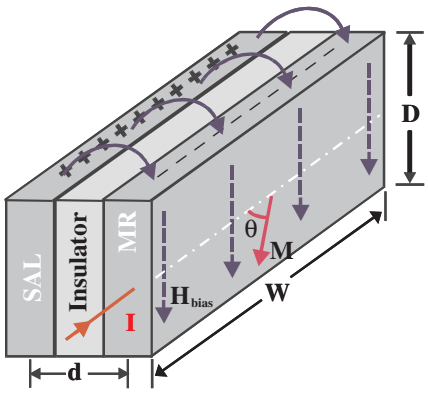


Figure 5. Schematic of an AMR sensor using an SAL bias. The sensor consists of an MR element as the active layer and a soft adjacent layer (SAL) for traverse bias. The MR element and the SAL are separated by an insulating spacer.

3.3. Giant Magnetoresistance

In the second half of the 1970s, researchers at IBM developed a technique which allows growing ultrathin films with extraordinary accuracy in thickness control. It was based on the vacuum evaporation technique, but with a several orders of magnitude lower base pressure as compared to the normal high-vacuum evaporator. At this base pressure and appropriate partial pressures of the source materials, it is possible to have a mean-free path of the evaporated materials that is larger than the distance between the source material and the substrate for most of the molecules or atoms. In other words, the evaporated molecules or atoms form beams, and impinge directly on the substrate surface to initiate the growth. Therefore, the technique is generally called molecular beam epitaxy (MBE) [62]. The word "expitaxy" was used because MBE had been mainly employed to fabricate semiconductor materials which are normally single-crystalline materials grown epitaxially on the substrates. The most attractive point of MBE is that it not only allows the growth of ultrathin single-layer film, but also makes it possible to grow layered structures of different materials. This opened an important field in materials science called artificial structures. In particular, when the layers are a few atomic layers thick, they form a so-called superlattice structure [63]. It also made it possible to create quantum-well structures in which the electrons of one type of material are confined by the potential steps formed between this material and the other types of materials at both sides. This led to tremendous advances in semiconductor devices. Being inspired by the work carried out by the semiconductor community, researchers in the surface-science community started to use the same technique to study surface magnetism and magnetic interactions across an ultrathin antiferromagnetic or nonmagnetic spacer [64]. These research activities led to the discovery of giant magnetoresistance in Fe/Cr superlattices [22]. As shown in Figure 6, the resistance of such a superlattice structure is high at zero field, decreases when a magnetic field is applied in both directions along the sample surface, and finally saturates at a field H_s of about 2 T. The MR ratio of superlattices with an [Fe(3 nm)/Cr(0.9 nm)] × 60 structure was measured to be about 45% at 4.2 K, which is much larger than the AMR effect [22]. It was proved soon afterward that the resistance of the superlattice structure depends on the relative orientation of the magnetization of the adjacent magnetic layers. At a Cr thickness of 0.9 nm, the magnetizations of adjacent Fe layers are antiferromagnetically coupled at zero field, but are realigned into the same direction when the external field is sufficiently high to overcome the exchange-coupling field.

The mechanism of GMR can be understood using the simple two-current model [65–67, 284, 285]. In this model,

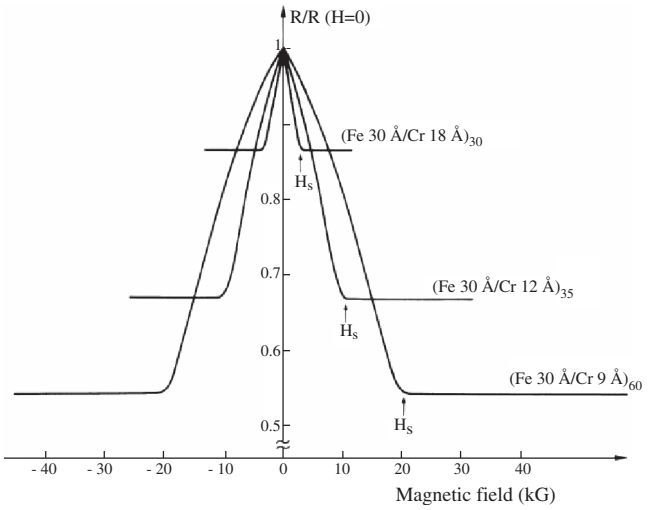


Figure 6. Magnetoresistance of Fe/Cr superlattices. Both the current and the applied field are along the same [110] axis in the film plane. Reprinted with permission from [22], M. N. Baibich et al., *Phys. Rev. Lett.* 61, 2472 (1988). © 1988, American Institute of Physics.

the electrical conduction of ferromagnetic materials is considered carried out by two independent channels of spin-up and spin-down electrons. Due to the different density of states distribution of spin-up and spin-down electrons near the Fermi level, the mean-free path for spin-up electrons is normally larger than that of spin-down electrons due to the larger scattering probability of s electrons to spin-down d-electron states (when the density of states of spin-down electrons is higher than that of the density of states of spin-up electrons). When all of the magnetic layers are ferromagnetically exchange coupled, the spin-up electrons will experience less scattering when they cross the nonmagnetic layer entering other magnetic layers within the spin-diffusion length, which is normally larger than the mean-free path. The resistivity of the spin-up electrons in this case can be considered as a constant over the film stack, and is denoted by $\rho_{\uparrow\uparrow}$. Similarly, the resistivity of the spin-down electrons can also be considered as a constant, and is denoted by $\rho_{\uparrow\downarrow}$. As in the simple two-current model, we ignore the spin-flip scattering, that is, no mixing of the two conduction channels; the total resistivity is given by $\rho_P = \rho_{\uparrow\uparrow}\rho_{\uparrow\downarrow}/(\rho_{\uparrow\uparrow} + \rho_{\uparrow\downarrow})$. However, the situation changes when the magnetic layers are antiferromagnetically coupled. This is because both the spin-up and spin-down electrons will experience frequent scattering when they cross the multiple layers. The two conduction channels will have the same resistivity, which is given by $(\rho_{\uparrow\uparrow} + \rho_{\uparrow\downarrow})/2$, leading to a total resistivity of $\rho_{AP} = (\rho_{\uparrow\uparrow} + \rho_{\uparrow\downarrow})/4$. The MR ratio is thus given by

$$\mathrm{MR} = \frac{\rho_P - \rho_{AP}}{\rho_{AP}} = -\left(\frac{\rho_{\uparrow\uparrow} - \rho_{\uparrow\downarrow}}{\rho_{\uparrow\uparrow} + \rho_{\uparrow\downarrow}}\right)^2 \tag{8}$$

The negative sign indicates that the resistivity at saturation state is lower than that at the zero-field state. A quantitative treatment of the GMR effect is possible using either the Boltzmann transport equation [68–70] or the quantum Kubo formula [71–73] using either the simple parabolic band structure or the more realistic band structures [74, 75]. These models can predict very well the dependence of GMR on the material and structural parameters of the multilayers. More details can be found in other review articles [76, 77]. For a magnetic/nonmagnetic multilayer with the film plane parallel to the xy plane, the linearized Boltzmann equation becomes [68–70]

$$\frac{\partial g^{\uparrow(\downarrow)}(z,\vec{v})}{\partial z} + \frac{g^{\uparrow(\downarrow)}(z,\vec{v})}{v_z \tau^{\uparrow(\downarrow)}} = \frac{eE}{mv_z}\frac{\partial f_o(\vec{v})}{\partial v_x} \tag{9}$$

where

$$g^{\uparrow(\downarrow)}(z,\vec{v}) = f^{\uparrow(\downarrow)}(z,\vec{v}) - f_0(\vec{v})$$

is the deviation of the electron distribution function from the equilibrium Fermi–Dirac distribution $f_0(\vec{v})$. Here, $\tau^{\uparrow(\downarrow)}$ represents the relaxation times for spin-up (spin-down) electrons, e is the elementary charge of electrons, m is the electron mass, $\vec{v}$ is the velocity, and E is the electrical field. The general solution of Eq. (9) is

$$g_{\pm}^{\uparrow(\downarrow)}(z,\vec{v}) = \frac{eE\tau^{\uparrow(\downarrow)}}{m}\frac{\partial f_0(\vec{v})}{\partial v_x}\left[1 + F_{\pm}^{\uparrow(\downarrow)}(\vec{v})\exp\left(\frac{\mp z}{\tau^{\uparrow(\downarrow)}|v_z|}\right)\right] \tag{10}$$

where $+(-)$ is for $v_z > 0$ ($v_z < 0$). $F_{\pm}^{\uparrow(\downarrow)}(\vec{v})$ can be obtained by using the boundary conditions for $g_{\pm}^{\uparrow(\downarrow)}(z,\vec{v})$ at both the interfaces and top and bottom surfaces. After $g_{\pm}^{\uparrow(\downarrow)}(z,\vec{v})$ is obtained, the total current per unit length in the y direction (assuming the field along the x direction) can be obtained by first integrating $(-e)v_x g_{\pm}^{\uparrow(\downarrow)}(z,\vec{v})$ in $\vec{v}$ space, followed by integrating with respect to z. The MR ratio can thus be obtained from the difference in conductivities between the parallel and antiparallel alignment of the magnetic layers.

3.4. Interlayer Coupling

The underlying mechanism of GMR is the long-range magnetic exchange coupling between transition metals across a nonmagnetic spacer, which was first observed in the Fe/Cr/Fe(001) system [78]. Although the GMR effect was originally observed in Fe/Cr superlattices with a fixed Cr layer thickness, subsequent studies have revealed the oscillatory behavior of interlayer coupling in this system and many others [79–81]. More importantly, these results were obtained not only from high-crystalline-quality samples grown by MBE, but also from polycrystalline samples deposited by sputtering. This has greatly accelerated the research in this field because of the wide availability of sputtering systems. As one of the typical results, Figure 7 shows the oscillatory behavior of magnetoresistance of Co/Cu superlattices observed by Parkin et al. on dc magnetron-sputtered samples [79]. Four well-defined oscillations with a period of about 1 nm are seen at both 4.2 and 300 K. This system later became the most important system for metal-based spintronic devices.

It was soon realized that the oscillation in GMR is a general phenomenon in ultrathin magnetic and nonmagnetic multiple layers, and its origin comes from the periodical switching from ferromagnetic to antiferromagnetic

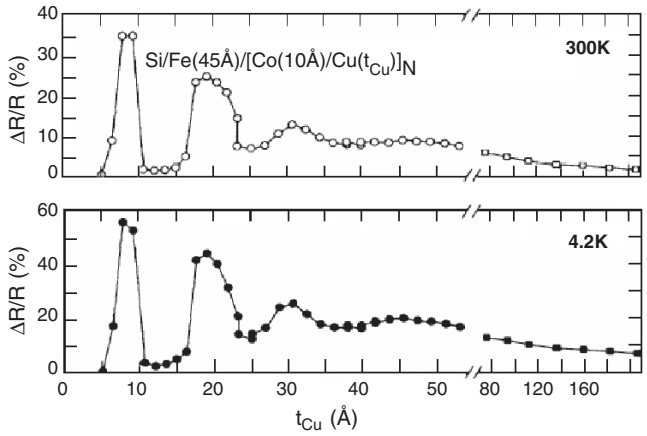


Figure 7. Dependence of saturation traverse magnetoresistance on Cu spacer layer thickness for a family of related superlattice structures of the form Si/Fe(4.5 nm)/[Co(1 nm)/Cu(t_{Cu})]$_N$. An additional Cu layer was deposited on each film structure such that the uppermost Cu layer was about 5.5 nm thick. The number of bilayers in the superlattice N is 16 for t_{Cu} below 5.5 nm (solid and empty circles) and 8 for t_{Cu} above 5.5 nm (solid and empty circles). Adapted with permission from [80], S. S. P. Parkin et al., *Phys. Rev. Lett.* 66, 2152 (1991). © 1991, American Institute of Physics.

alignment and vice versa of the adjacent ferromagnetic layers when the thickness of the nonmagnetic layer varies every few atomic layers [82]. It was found that both the GMR and the saturation magnetic field of Fe/Cr multiple layers oscillate at a period of about 1.5 nm. The latter suggests that the relative angle of the magnetization direction of the two ferromagnetic layers oscillates between 0 and 180° because it is easier to saturate the two layers when they are in parallel alignment, and difficult when they are in antiparallel alignment. This was soon confirmed by the neutron [83] and Brillouin scattering experiments [84]. Furthermore, detailed investigations have revealed the existence of two coupling periods [85–90].

Various models have been proposed to account for the oscillatory behavior of exchange coupling of the magnetic/nonmagnetic multiple layers, which include primarily the quantum-well model [91–97] and the Ruderman–Kittel–Kasuya–Yoshida (RKKY) interaction model [98–101]. Both models have been able to describe the origin of the exchange coupling, in particular, the oscillations in the sign and strength of the coupling. In what follows, we explain briefly the RKKY model by closely following the treatment by Bruno and Chappert [100, 101] and Fert and Bruno [102]. The interaction between two ferromagnetic layers across a metallic spacer has its basis in the interaction between two magnetic impurities in a host metallic material. According to the RKKY model, the effective exchange interaction between two spins $\vec{S}_i$ and $\vec{S}_j$ is given by

$$H_{ij} = J(\vec{R}_i - \vec{R}_j)\vec{S}_i \cdot \vec{S}_j \tag{11}$$

where the exchange integral is

$$J(\vec{R}_i - \vec{R}_j) = -\frac{1}{2}\left(\frac{A}{V_0}\right)^2 \frac{V_0}{(2\pi)^3}\int d^3\vec{q}\chi(\vec{q})\,\exp[i\vec{q}\cdot(\vec{R}_i - \vec{R}_j)] \tag{12}$$

Here, V_0 is the atomic volume and

$$\chi(\vec{q}) = 2\mu_B^2 \frac{V_0}{(2\pi)^3}\int d^3\vec{k}\frac{f(\varepsilon_{\vec{k}}) - f(\varepsilon_{\vec{k}+\vec{q}+\vec{G}})}{\varepsilon_{\vec{k}+\vec{q}+\vec{G}} - \varepsilon_{\vec{k}}} \tag{13}$$

is the susceptibility of the host materials. The integrations over $\vec{q}$ and $\vec{k}$ are to be performed within the first Brillouin zone, and $\vec{G}$ is a vector such that $\vec{k} + \vec{q} + \vec{G}$ belongs to the first Brillouin zone. In the case of free electrons at zero temperature, the Fermi surface is a spherical surface with a radius defined by the Fermi wavevector $\vec{k}_F$. The probability distribution function $f(\varepsilon_{\vec{k}})$ drops to zero sharply at the Fermi surface, resulting in a logarithmic singularity at $\vec{q} = 2\vec{k}_F$ for $\chi(\vec{q})$. The Fourier transforms of $\chi(\vec{q})$ give oscillations to J with a period of $\Lambda = \lambda_F/2$, and λ_F is the Fermi wavelength.

For interactions between two ferromagnetic layers (hereafter, we refer them to as FM_1 and FM_2), the interaction integral may be obtained by summing over all of the pairs of spins belonging to FM_1 and FM_2. For the sake of simplicity, we assume that both layers are one atomic layer thick. Without the loss of generality, we first look at the interaction between a single spin located at position O of FM_1 and all of the spins belonging to FM_2, which is given by

$$I_{1O,2} = \frac{d}{V_0}S^2 \sum_{j\in FM2} J(\vec{R}_O - \vec{R}_j) \tag{14}$$

The above summation can be further simplified by treating FM_2 as a thin layer with a uniform spin distribution, that is,

$$\sum_{F2} \to \frac{d}{V_0}\int_{F_2} d^2\vec{R}_{||} \tag{15}$$

Here, $\vec{R}_{||}$ is the in-plane component of $(\vec{R}_O - \vec{R}_j)$. The interaction integral is then given by

$$I_{1O,2}(z) \approx \left(\frac{A}{V_0}\right)^2 S^2 \frac{m}{16\pi\hbar^2}\frac{d^2}{z^2}\sin(2k_F z), \text{ for } z \to \infty \tag{16}$$

which is also an oscillation function with a period of $\Lambda = \lambda_F/2$. This is too short to explain the long oscillation periods observed experimentally. Subsequent studies pointed out that the contradiction comes from the invalid assumption that the thin FM_2 layer has a continuous and uniform spin distribution. By taking into account the discreteness of the spin distribution in the FM_2 plane, the interaction integral is given by

$$I_{1O,2}(z) = -\frac{1}{2}\left(\frac{A}{V_0}\right)^2 S^2 \frac{d}{(2\pi)^3}\int_{-\infty}^{+\infty} dq_z \times \exp(iq_z z)\int d^2\vec{q}_{||}\chi(\vec{q}_{||}, q_z) \sum_{\vec{R}_{||}\in FM_2} \exp(i\vec{q}_{||}\cdot\vec{R}_{||}) \tag{17}$$

The last term suggests that $\vec{q}_{||}$ cannot take any arbitrary value but the in-plane reciprocal lattice vector $\vec{G}_{||}$; otherwise, its value will be zero. This gives an effective wavevector

$[(2k_F)^2 - G_{||}^2]^{1/2}$, which is responsible for the multiperiodicity due to different $G_{||}$ in different directions. Figure 8 shows the cross section of the Fermi surface for an fcc (001) spacer. The first Brillouin zone (FBZ) and the auxiliary prismatic zone are represented by the solid and dashed contours. The wavevectors giving rise to oscillatory interlayer coupling are indicated by the horizontal bold arrows: (a) free-electron approximation, and (b) Cu Fermi surfaces. Also shown are the vectors giving rise to oscillations in the (111) direction (oblique arrows).

On the other hand, the long oscillation period can also be understood based on the argument that the spacer thickness does not vary continuously in Eq. (16); rather, it only takes discrete values $z = (N+1)a$ with N an integer number [99, 103, 104]. Because of this discrete sampling or aliasing effect, one obtains an effective period

$$\Lambda = \frac{\pi}{|k_F - n\pi/a|} \tag{18}$$

where n is an integer chosen such that $\Lambda > 2a$, with a being the atomic spacing of the spacer in the thickness direction. This gives a period which is much larger than $\lambda_F/2$ (see Fig. 9).

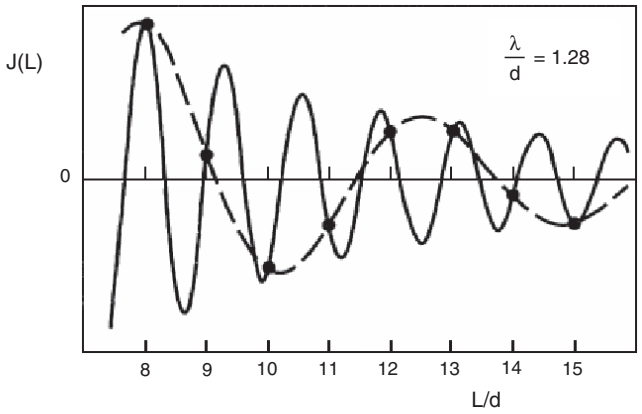


Figure 9. Full curve: exchange-coupling strength of two ferromagnetic layers across a monovalent fcc (100) metal (arbitrary units) calculated using the continuum version of the RKKY model. Broken curve: the actual coupling strength with the experimentally measured periodicity for $L = Nd$. Here, L is the thickness of the spacer, d is the atomic layer spacing in the thickness direction, and N is an integer number. Adapted with permission from [99], R. Coehoorn, *Phys. Rev. B* 44, 9331 (1991). © 1991, American Institute of Physics.

3.5. Spin Valve

The research on read sensors using the GMR effect commenced soon after the GMR effect was reported. This is because it gives a much higher MR ratio as compared to its AMR counterpart. However, the original GMR structure consisting of Fe/Cr, Co/Cu, or other types of magnetic multiple layers could not be applied to read sensors in the original form because of their large saturation fields, and also the nonlinear response near the zero-field point. The most straightforward way to reduce the saturation field is to reduce the exchange coupling between the ferromagnetic layers through increasing the thickness of the spacer. However, a completely decoupled or weakly coupled multilayer structure does not function properly as a sensor because of the possible inconsistent movement of all of the magnetic layers under an external field. One of the possible ways to overcome the drawback of the decoupled GMR stack is to use two decoupled layers with one of them much softer than the other, so that the former will respond to a small field, while the latter will only change its magnetic state when it is subjected to a large field [105, 106]. This kind of structure, in principle, can form a good sensor, provided that the hard layer is not only truly hard, but also has the same crystalline structure as that of the soft layer. Unfortunately, this kind of material is still not available at present. Therefore, most of the experiments on decoupled bilayers have been performed on NiFe/Cu/Co or Co/Cu/Co systems. In the latter case, the two Co layers may have different coercivities, and their values are dependent on the individual layers' thicknesses. Although a relatively large magnetoresistance with a small saturation field has been obtained in these kinds of structures, they were still unsuitable for data storage applications due to the instability of the hard layer. However, this type of structure has served as a good object for studying spin-dependent transport in magnetic/nonmagnetic multilayers [107, 108]. They can also be used for MRAMs [109]. The magnetic stiffness of the hard layer may be improved by using an antiferromagnetically coupled hard/Ru/soft triple-layer structure to replace the single-layer structure [110, 111].

To address the issues involving noncoupled magnetic multilayers, IBM invented a device called the spin valve in 1991 [24, 112, 113]. A typical spin valve in its simplest form

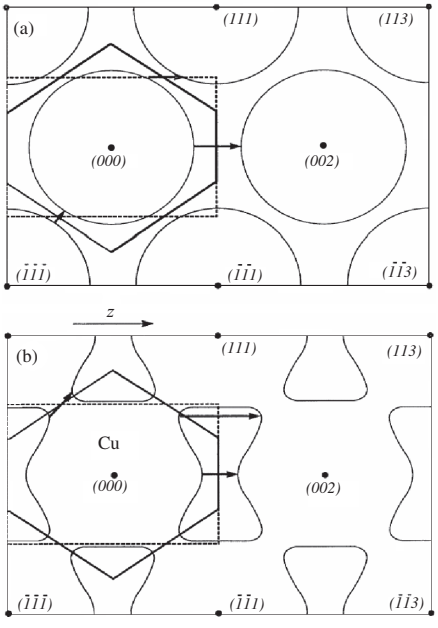


Figure 8. Cross section of the Fermi surface for an fcc (001) metal and the corresponding vectors giving rise to oscillatory interlayer coupling in both the (001) and (111) directions (indicated by the horizontal and oblique arrows, respectively). (a) Free-electron model. (b) Cu Fermi surface. Adapted with permission from [100], P. Bruno and C. Chappert, *Phys. Rev. Lett.* 67, 2592 (1991). © 1991, American Institute of Physics.

consists of two ferromagnetic layers separated by a nonmagnetic spacer and an antiferromagnetic (AFM) layer in contact with one of the ferromagnetic layers (see Fig. 10). The thickness of the spacer is chosen such that there is little exchange coupling between the two FM layers. The magnetization of one of the FM layers which is in direct contact with the AFM layer is "pinned" by the latter, and thus this FM layer is commonly called a pinned layer. On the other hand, the magnetization of the other FM layer is free to rotate to respond to an external field, and thus it is called a free layer. The typical material for the FM layer is NiFe or CoFe, while that for the spacer is generally copper. A thin layer of Co or CoFe with high Co composition is often added at the FM and Cu interface so as to increase the MR ratio due to its high polarization ratio and less interdiffusion at the interface with Cu [114, 115]. The choice of AFM is an issue of high complexity due to many practical considerations, which will be discussed shortly.

When being used as a read sensor [116], a constant current is applied to the sensor, and its voltage change caused by the fringe-field variation from the media is detected as the readout signal. In order to increase the linear density, usually two shielding layers are used to "shut off" the influence of neighboring bits (see Fig. 11). Depending on how the current is applied, there are two different types of sensors, that is, the Current-in-plane and current-perpendicular-to-plane sensors, as discussed earlier. As the name suggests, in the former case, the current flows in the plane of the sensor stack (see Fig. 11), while in the latter case, the current flows perpendicular to the plane. The small resistance of the metallic multilayers makes it quite challenging to measure the magnetoresistance of the CPP structure. Several approaches have been taken to address this issue, including the use of superconducting electrodes and reduction of the lateral size using lithography [117, 118] or selective growth in the pores of nanoporous membranes [119–121]. The results obtained showed that the CPP MR ratio is much larger than the CIP MR ratio. Alternatively, the GMR can also be measured with the current-at-an-angle (CAP) through growing the GMR structure on grooved substrates [122]. Since, so far, CIP sensors dominate in practical applications, we discuss the CIP sensors first.

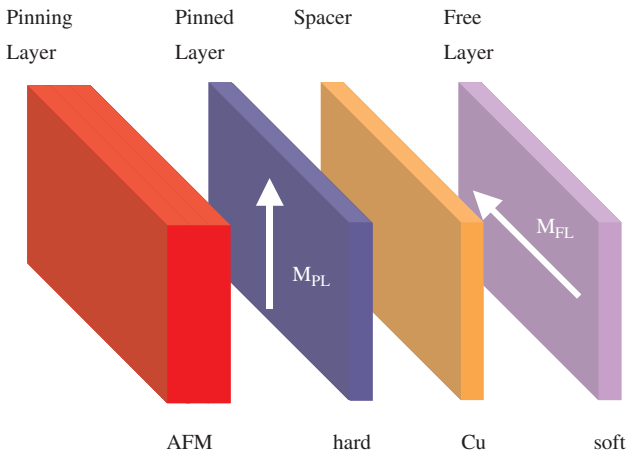


Figure 10. Schematic of a simple spin-valve structure.

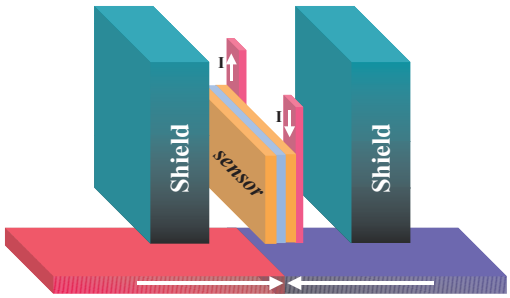


Figure 11. Spin valve used as a read sensor in CIP mode.

The M–H and MR–H curves of a typical spin valve are shown schematically in Figure 12. Due to the small thickness of the spacer, there always exists an interlayer coupling between the free and pinned layers. The nature of the interlayer coupling can either be ferromagnetic or antiferromagnetic, depending on whether the orange peel or RKKY coupling dominates. We will come back to this topic shortly. For the particular spin valve whose characteristics are shown schematically in Figure 12, the free and pinned layers are ferromagnetically coupled. Therefore, the magnetization of the spin valve is at maximum, while the resistance is at minimum at zero field. When the applied field increases in the opposite direction of the pinning field (in this case, in the positive direction), the magnetization of the free layer was first reversed, resulting in an antiparallel alignment of the free and pinned layers. This corresponds to the high-resistance state. When the applied is further increased, it will eventually reverse the magnetization of the pinned layer. This will realign the magnetization of the pinned layer with that of the free layer, recovering the low-resistance state. Decreasing of the applied field after the saturation in the positive direction will cause the magnetization of the pinned layer to be reversed first, followed by the reversal of the free layer, and finally, the two magnetic layers realign with each other again at negative fields. In general, a large exchange bias field and a small interlayer-coupling field are desirable for read-head applications.

The performance of a read sensor is normally characterized by the size of the MR effect. It can be either the MR ratio or the sheet resistance change ΔR_S. The former is given by [24, 123–125]

$$\frac{\Delta R_S}{R_S} = \left(\frac{\Delta R_S}{R_S}\right)_0 \left(\frac{1-\cos(\theta_F - \theta_P)}{2}\right) \tag{19}$$

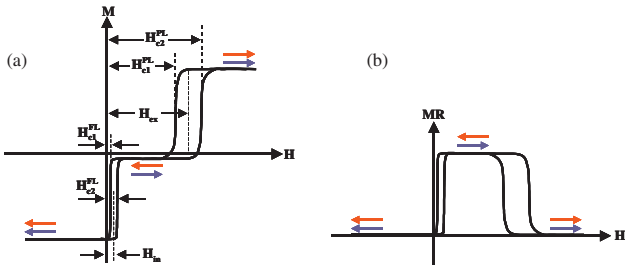


Figure 12. Schematics of M–H (a) and MR–H (b) curves of a typical spin valve (H_{ex}: exchange-bias field; H_{in}: interlayer coupling field between free and pinned layer; $(H_{c1}^{FL} - H_{c2}^{FL})$: coercivity of free layer; $(H_{c1}^{PL} - H_{c2}^{PL})$: coercivity of pinned layer).

where R_S is the sheet resistance and ΔR_S the change in sheet resistance. $(\Delta R_S/R_S)_0$ is the maximum MR ratio when the magnetizations of free and pinned layers are switched from parallel alignment to antiparallel alignment. θ_F and θ_P are the angles between the parallel direction and the magnetization direction of the free and pinned layers, respectively. It is apparent from this equation that a linear response can be obtained through setting θ_P to 90°. This is because $\sin\theta_F$ is proportional to the external field, as we discussed in Section 3.2 for AMR sensors. In addition to its linear response, the spin valve also exhibits a much higher MR ratio as compared to the AMR sensors. However, even though that is the case, there has been and there is still a strong demand for the improvement of the spin-valve sensors. Figure 13 shows the estimated track width and required sensor sensitivity at an aspect ratio of 4:1 for different areal densities by assuming that the sensor has to generate an output voltage of 1 mV. The solid and dashed curves are obtained by assuming a read track width of 80 and 60% of the track pitch, respectively. From this figure, one can see that a threefold increase of the sensor sensitivity is required for increasing the areal density from 100 Gbits/in^2 to 1 Tbits/in^2.

From Eq. (19), one can see that the MR ratio of a spin valve is affected by two factors of different origin. The prefactor of the right-hand-side term of Eq. (19), that is, $(\Delta R_S/R_S)_0$, is mainly determined by the spin-dependent electronic transport properties of the spin valve, while the rest is more or less determined by its magnetic properties. As summarized in Figure 14, the improvement in spin-valve design has proceeded from both the magnetic and electronic aspects. In a certain sense, the original spin valve proposed by IBM can be regarded as the first step in the improvement of the pseudospin valve that still remains the simplest, yet the most attractive design should the materials problem be solved in the future. In what follows, we first discuss the

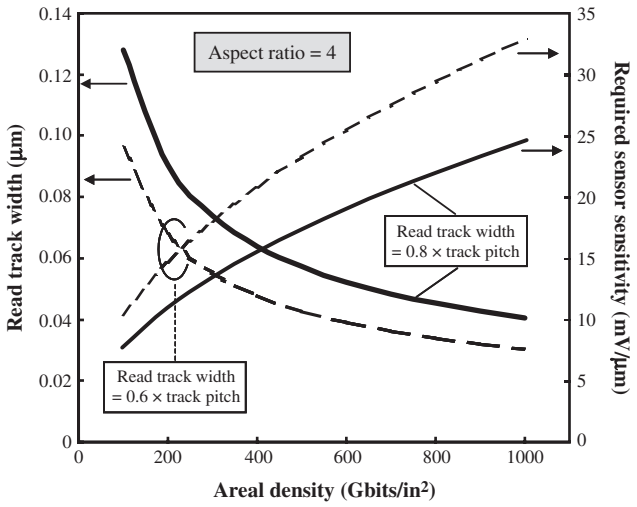


Figure 13. Estimated read track width and required sensor sensitivities at different areal densities by assuming an aspect ratio of 4:1. The solid and dashed curves are obtained by assuming the read track width to be 80 and 60% of the track pitch, respectively. The sensitivity requirement is obtained by assuming that the sensor would generate an output voltage of 1 mV.

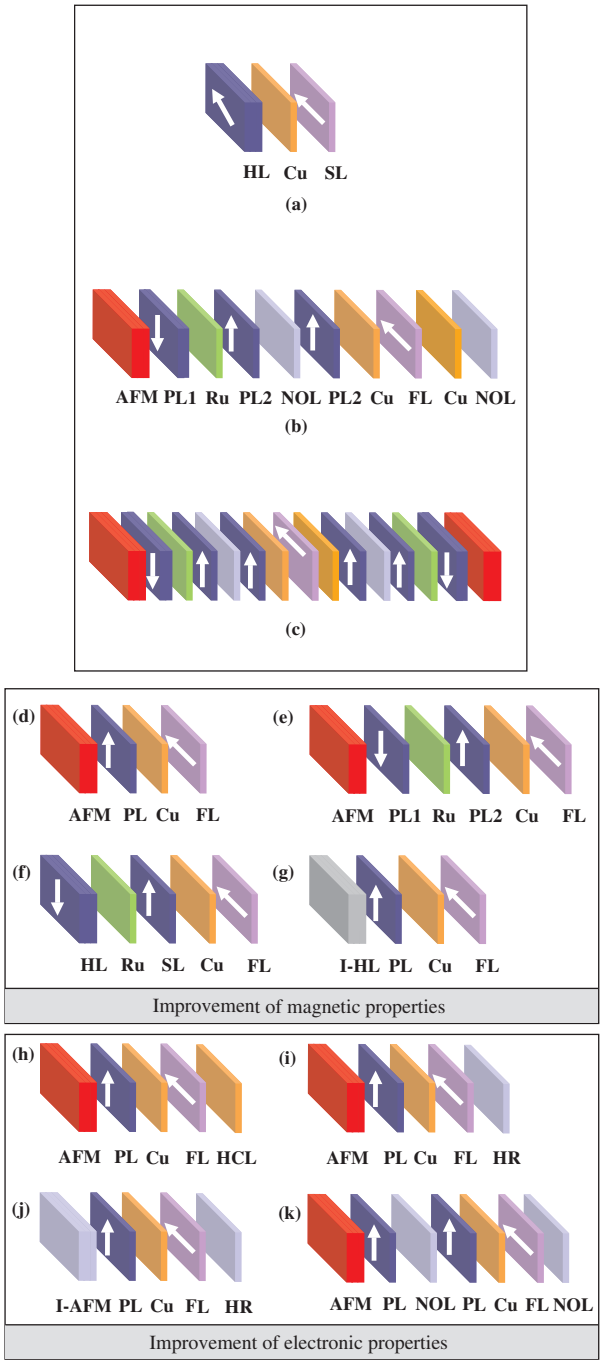


Figure 14. Evolution of spin valves. (a) Pseudospin valve. (b) Advanced single spin valve. (c) Advanced dual spin valve. (b) and (c) are the results of continuous improvements in both the magnetic (d)–(g) and electronic (h)–(k) designs of the pseudospin valves. (d) Original spin-valve invented by IBM. (e) Synthetic spin valve. (f) Synthetic spin valve using a hard magnet. (g) Exchange spring spin valve using an insulating hard magnet. (h) Spin-filter spin valve using a back layer or a high-conductance layer. (i) Specular spin valve. (j) Specular spin valve using an insulation AFM. (k) Specular spin valve using nano-oxides. The acronyms used are: HL—hard layer, SL—soft layer, PL—pinned layer, FL—free layer, PL1 (2)—first (second) portion of the pinned layer, I-HL—insulating hard layer, HCL—high-conductance layer, HR—high-specularity reflective layer, NOL—nanooxide layer, AFM—antiferromagnetic layer, I-AFM—insulating antiferromagnetic layer.

magnetic aspect of spin valves, followed by a discussion on the electronic aspect. We will refer back to Figure 14 again when necessary.

4. MAGNETIC ASPECTS OF SPIN VALVES

Like a transistor which needs a proper bias to work properly, the spin valve also needs a proper magnetic bias so as to achieve a linear response with minimized asymmetry, if any [13, 15, 55, 123]. The main tasks of magnetic design of spin valves are: (1) to set the magnetization of the pinned layer as rigid as possible in the vertical direction (see Fig. 10), (2) to make the free layer as soft as possible, and to align its initial magnetization direction parallel to the media surface, (3) to eliminate Baukhausen noise induced by domain wall movement, and (4) to make the sensor thermally stable at elevated temperature under normal operation conditions and corrosion resistive. Of course, in practical cases, one also has to consider the ease of manufacturing and reliability of the devices. To achieve all of these goals, one needs to set a proper traverse and longitudinal bias for the spin valve using appropriate materials. The former is employed to "pin" the magnetization of the pinned layer as rigid as possible, while the latter's role is to maintain the single-domain status of the free layer, and at the same time, without sacrificing its softness.

4.1. Traverse Bias

4.1.1. Exchange Bias

Traverse bias is achieved through the use of exchange bias (EB) formed at FM/AFM interfaces [Fig. 10 and Fig. 14(d)], which was discovered about half a century ago [126]. The exchange bias at the FM/AFM interface is usually established through a magnetic annealing and cooling process, in which the sample is first heated up to above the Néel temperature of the AFM, and then is cooled down to room temperature in the presence of a magnetic field with a strength of one to a few tesla. The exchange-biased FM/AFM bilayer is characterized by a shift of the hysteresis loop in the opposite direction of the exchange-bias field direction and an increase of the coercivity of the FM layer, as shown schematically in Figure 15(a).

The exchange bias in AFM/FM bilayers is a very complex issue. Despite intensive studies, the phenomenon itself is still not fully understood [127–130]. A number of theoretical models have been proposed to account for the experimental data, which include, but are not limited to: (1) the coherent rotation model [131], (2) the Néel model [132], (3) the random interface model, (4) the AFM domain wall model [133, 134], (5) the interface spin canting model [135], (6) the random and compensated interface model, and so on [136–139]. Obviously, the coherent rotation model is the simplest one which assumes that: (1) both the FM and AFM layers are single-crystalline materials, (2) the AFM layer has a rigid spin configuration which is not affected by any strength of an external field, and its interfacial layer is fully uncompensated, and (3) the magnetization reversal of the FM layer is accomplished through coherent magnetization

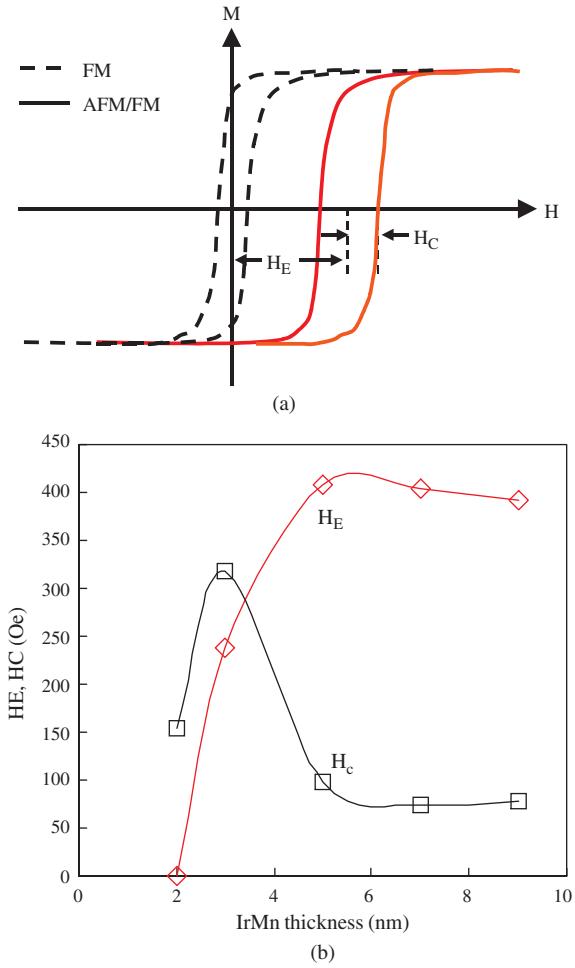


Figure 15. (a) Schematic hysteresis curves of a single FM layer (dotted line) and an FM/AFM exchange-coupled bilayer (solid line). (b) Dependence of exchange-bias field on the thickness of AFM material for an NiFe (3 nm)/IrMn bilayer system.

rotation. Under these assumptions, the exchange field is given by

$$H_{EX} = \frac{2J\vec{S}_{FM} \cdot \vec{S}_{AFM}}{a^2 M_{FM} t_{FM}} \tag{20}$$

where J is the exchange parameter, $\vec{S}_{FM}(\vec{S}_{AFM})$ is the spin of the interfacial atom of the FM (AFM) layer, $M_{FM}(t_{FM})$ is the magnetic moment (thickness) of the FM layer, and a is the cubic lattice constant. Although the above equation does agree with the experimental observation that H_{EX} decreases with t_{FM}, it yields a value for H_{EX} that is by orders of magnitude larger than that of the experimental data. The simple model fails to explain the asymmetry observed in the hysteresis either. This has prompted the proposal of various improved models. Among them, two models have been quite successful in explaining the observed value of H_{EX}. Mauri et al. proposed a model that is able to explain the observed exchange bias without the need to remove the condition of strong interfacial coupling at the FM/AFM interface [134]. This is made possible by allowing the formation of parallel domain walls, either in the AFM layer or in the FM layer,

wherever the energy is lower. By assuming that the domain wall forms at the AFM side of the interface, they obtained the following expression for the exchange field:

$$H_{EX} = \frac{2\sqrt{A_{AF}K_{AF}}}{M_{FM}t_{FM}} \tag{21}$$

where $A_{AF}(K_{AF})$ is the exchange stiffness (magnetocrystalline anisotropy energy constant) of the AFM layer.

Malozemoff argued that any realistic model should not be based on the assumption of an atomically perfect uncompensated AFM surface, and proposed a random interface model [133]. It was shown that, for a real interface with finite roughness, it becomes energetically favorable for the antiferromagnet to break up into domains with perpendicular domain walls. This model yields an exchange-coupling field of

$$H_{EX} = \frac{2z\sqrt{A_{AF}K_{AF}}}{\pi^2 M_{FM}t_{FM}} \tag{22}$$

where z is the number of order unity. Therefore, the exchange field predicted by both the AFM domain wall model and the random interface model is of the same order; this is because both are determined by the characteristic domain wall energy.

The parallel domain wall model requires the existence of a spiral spin structure in the AFM layer [140], which has been observed recently in both the conventional FM/AFM bilayers [141] and artificial FM/AFM bilayers [142]. The formation of an AFM domain wall is found to be affected by the magnetization of the FM layer. The exchange coupling between the FM and AFM layers occurs on a domain-by-domain basis [143]. Due to the dominant role of the AFM domain walls in exchange coupling, the exchange bias is generally sensitive to the defects in the AFM layer.

4.1.2. AFM Materials

Above is a far too brief introduction of the exchange-coupling mechanism at the FM/AFM interfaces. An in-depth discussion of this topic can be found in several excellent review papers [127–130]. In what follows, we turn to different material systems that have been developed for practical applications in spin valves [144, 145]. As discussed earlier, the key to the proper operation of a spin valve is to "pin" the magnetization of the reference layer in a direction that is perpendicular to the media surface, which forms a 90° angle with respect to the easy axis of the free layer at zero field. The former is achieved using the exchange bias at the AF/AFM interface. To realize a reliable and linear operation of the spin valve, the following issues have to be considered carefully when choosing the AFM materials.

1. Size of the exchange bias field: As a subtle rotation of the magnetization of the pinned layer will cause asymmetry and degradation of the output signal, the pinning field must be sufficiently high (>300 Oe) so as to make the reference layer "rigid" under a moderate external field.
2. Thermal stability: The exchange-bias field usually decreases with temperature. In what way and how fast it decreases, however, depends not only on the intrinsic properties of the materials, but also their structural properties, thickness, and techniques that are used to prepare the materials. The temperature at which the exchange-bias field vanishes is called the blocking temperature. A high blocking temperature (>250 °C) is required to ensure that the sensor functions properly not only at room temperature, but also at elevated temperatures. This is important because the temperature of the active region of the sensor may rise to 60–120 °C during normal operation.
3. Large resistivity of the AFM layer: As the AFM layer itself does not contribute to the MR effect, a high-resistivity material is desirable for suppressing the current-shunting effect. In this sense, the ideal AFM material should be an insulator. Unfortunately, most of the oxide-based AFM materials like NiO can hardly satisfy other requirements, and thus are unsuitable as AFM materials for practical device applications.
4. Small critical thickness: The exchange-bias field is found to be quite insensitive to the thickness of the AFM when it is sufficiently thick. However, it starts to drop rapidly when the thickness decreases to a certain value, which is largely determined by the properties of the AFM material. As an example, Figure 15(b) shows the results for an $Ni_{80}Fe_{20}$(3 nm)/IrMn bilayer system. The onset thickness for this process is called the critical thickness. Considering the temperature dependence of the exchange field, the thickness of the AFM must be chosen well beyond the critical thickness in practical devices. However, a large AFM thickness will result in a large total thickness of the sensor, which is undesirable for sensors for high-areal-density hard disk drives. Therefore, AFM materials with smaller critical thickness are desirable for spin-valve applications.
5. Good corrosion resistance: AFM materials with poor corrosion resistance are unsuitable for practical applications.
6. Low-temperature process: The thermal annealing process may affect the performance of spin valves if the annealing temperature is too high. Therefore, a low-temperature process is preferred.

The AFM materials investigated so far include: (1) Mn-based alloys, (2) Cr-based alloys, and (3) Fe, Co, or Ni-based oxides [144, 145]. Among them, Mn-based alloys are the most widely studied, and several of them have already been applied to practical spin-valve sensors. The Mn-based alloys can be roughly divided into two groups. One group includes FeMn, IrMn, RhMn, RuMn, and the other includes NiMn, PtMn, PdMn, and some of the ternary alloys of these elements. The crystalline structure of the first group is fcc, while that of the second group is fct (CuAu–I). FeMn/NiFe (CoFe) is the most widely studied system which exhibits an exchange bias of about 420 Oe for a 4 nm thick NiFe without thermal annealing. However, it has a low blocking temperature of about 150 °C and poor corrosion resistance, and thus is unsuitable for read sensor applications. Another system that does not require thermal annealing is IrMn/CoFe

(or NiFe). It has a better corrosion resistance and smaller critical thickness as compared to FeMn. Both are desirable for sensor applications. However, its moderate blocking temperature of 250 °C limits its applications in a real drive environment. Although, in general, a thermal annealing process (at 200–300 °C) is required to establish the AFM phase, the (Ni, Pt, Pd) Mn family gives a much higher blocking temperature (350–450 °C), better corrosion resistance, and a higher exchange-bias field (500–800 Oe), and thus is being used in real devices. The primary drawback of this group of materials is their large critical thickness, which may limit their applications in ultrahigh-density recording applications. Table 1 lists the properties of some widely studied AFM materials [144–160].

4.1.3. Synthetic AFM

As discussed earlier, the exchange-bias field is inversely proportional to the thickness of the FM layer. For practical applications, however, one needs a certain thickness for the pinned layer so as to optimize the electrical properties of the sensor. This affects the sensor in two aspects. One is the reduction in the exchange-bias field, and the other is the magnetostatic coupling with the free layer. Although one can use the sense current field to counterbalance the magnetostatic field, it may sacrifice the stability of the pinned layer. A novel structure which can solve both problems is the synthetic ferrimagnet which consists of two antiferromagnetically coupled FM layers via a thin Ru layer [see Fig. 16(a)]. The corresponding synthetic antiferromagnet in which one of the FM layers of the synthetic ferrimagnet is exchange coupled to an AFM layer can be used as the pinning and pinned layers for spin valves [161–171].

The partial cancellation of the magnetic moments from both layers gives a smaller net moment, and thus an increased exchange-bias field and a reduced magnetostatic-coupling field. In addition to these advantages, it also provides a better thermal stability, although the blocking temperature itself remains the same. Shown in Figure 16(b) is a schematic drawing of a typical spin valve using the synthetic AFM as the pinned layer [also see Fig. 14(e)]. As the exchange bias is inversely proportional to the magnetic moment thickness product of the pinned layer, compared to parallel alignment, the exchange-bias field in antiparallel alignment is enhanced by a factor of [21]

$$f_{P-AP} = \frac{M_{1s}t_1 + M_{2s}t_2}{M_{1s}t_1 - M_{2s}t_2} \tag{23}$$

where $M_{1s}(M_{2s})$ and $t_1(t_2)$ are the magnetic moment and thickness of the first (second) pinned layer, respectively. On the other hand, the demagnetizing field is proportional to the magnetic moment thickness product of the pinned layer itself. Therefore, the stability of the pinned layer is effectively increased by a factor of

$$[(M_{1s}t_1 + M_{2s}t_2)/(M_{1s}t_1 - M_{2s}t_2)]^2$$

Although it is not shown in Figure 16, the synthetic ferrimagnet can also be used as the free layer, which can effectively reduce the magnetic thickness of the free layer without sacrificing its physical thickness [172]. However, the longitudinal bias may affect the stability of the synthetic layer [173].

Table 1. Properties of widely studied AFM materials.

Materials	H_{ex} (Oe) (NiFe 4 nm[a])	Required thickness (nm)	T_B (°C)	Resistivity ($\mu\Omega \cdot$ cm)	Corrosion resistance	Annealing
FeMn	420	7[b]	150[b, c], 165[e], 175[f], 170[j]	150, 130[b]	Poor	No[b]
IrMn	200	7[d], 7.5[e], 10[l]	190/230[d], ~250[e, f], 130[l]	200, 325[d]	Good[a, k] Moderate[e]	No[k]
NiMn	>390	>25[b], 35[d], 30[e]	450[b], >425[d], 430[g], 400[f, j]	175[b], 210[d]	Moderate Good[e, k]	Required[b]
PtMn	500	30	380[p], 400[f], 310[j]	~200	Good	Required
PdPtMn	280	25[d, m]	300[m], 350[d, f]	185[d]	Good	Required
CrMnPt	220	30[a, o], 35[d]	380[a, o], 300[d]	310–345[a], 300–350[o] 360[d]	Good	No
NiO	200	35[b], 30[e], 50[h]	200[b, h, i], 190[e]	Insulator, ≥10[8b]	Good Excellent[e]	No[b]
α–Fe_2O_3	40–75	100	200–250, 250[q], 280–300[r]	Insulator	Good	No

Note: [a][144], [b][146], [c][147], [d][145], [e][148], [f][149], [g][150], [h][151], [i][152], [j][153], [k][154], [l][155], [m][156], [o][157], [p][158], [q][159], [r][160].

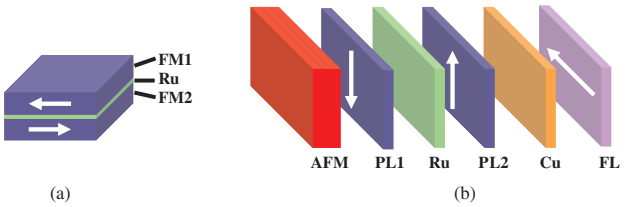


Figure 16. Schematic of a synthetic ferrimagnet (a) and spin valve using the synthetic ferrimagnet as a pinned layer (b).

4.1.4. Hard/Ru/Soft Synthetic Ferrimagnet

The use of AFM as the pinning layer has two drawbacks: (1) all of the useful AFMs reported so far for spin valves are metals which contribute negatively to the MR due to the current-shunting effect, and (2) most AFMs have a moderate or low blocking temperature, and some of them also show poor corrosion resistance. Based on this background, there have been several works in which hard/Ru/soft synthetic ferromagnets have been used as the pinning and pinned layers [see Fig. 14(f)] [110, 111, 174]. Preliminary experimental results showed that these types of spin valves have a good thermal stability [110, 111]. However, it would be quite challenging to improve other properties of the spin valves due to either the mismatch of the crystalline structure of the hard and soft layer or the insufficient hardness of the hard layer.

4.1.5. Exchange Spring

Further simplification of the synthetic ferrimagnet would be the hard/soft exchange spring structure [Fig. 14(g)]. This type of structure was proposed recently by Carey et al. using cobalt ferrite as the pinning layer [175]. A high MR ratio, 12.8%, and a high pinning field, 1500 Oe, have been obtained without compromising the softness of the free layer. In this particular case, the pinning layer is an insulator which may also help to reduce the shunting current, and thus is desirable for obtaining a high MR ratio.

4.2. Longitudinal Bias

The traverse bias is to set a proper direction for the magnetization of the pinned layer. For a spin valve to function properly, one also needs to set a proper longitudinal bias for the free layer. This is to both obtain a good linearity with less asymmetry and to suppress the Baukhausen noise. The former is achieved through setting the magnetization of the free layer at 90° with that of the pinned layer, parallel to the media surface. This, in turn, can be achieved through first inducing an easy axis in an appropriate direction during deposition of the free layer, and then using the shape anisotropy to stabilize it in the same direction. However, the shape anisotropy alone may become insufficient as the aspect ratio of the sensor decreases. The following additional magnetic fields need to be taken into account when designing the sensor: (1) the fringe field from the pinned layer, (2) the RKKY and orange-peel coupling field from the pinned layer, or so-called interlayer coupling field, (3) the current-induced field from highly conductive layers such as the spacer layer, and (4) other fields such as imaging fields from different sources [123]. Figure 17 shows schematically

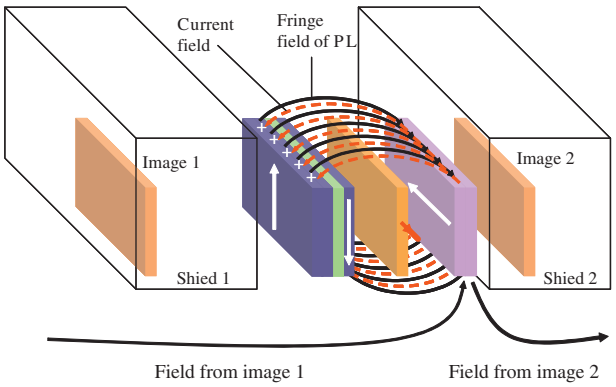


Figure 17. Different kinds of magnetic fields exerted on the free layer. For a spin valve to function properly, it is of importance to balance out all of the components so as to make the net field as close to zero as possible. Here, PL (FL) stands for pinned layer (free layer).

the flux lines of various field components exerted on the free layer. It is worth noting that the direction and magnitude of the current-induced fields are adjustable through controlling the sensing current, whereas the various field components from the pinned layer are fixed once the sensor fabrication is completed. Therefore, in practice, it is of importance to ensure that magnetic fields from sources other than the sense current be able to be compensated later by the current field under normal operation conditions. Ideally, all of the external fields other than the signal field from the media should be reduced to zero so as to obtain a high sensitivity, good linearity, and null asymmetry for the read sensor. However, this also means that the sensor is too susceptible to external disturbances. This will induce noises or baseline popping and a shift in the readout signal, in particular, the domain formation and movement-induced Baukhausen noise. The latter is an issue of high complexity because it depends on many factors, such as the material and shape of the free layer and the process to form it and the effect of other layers. Therefore, as in the case of AMR sensors, a longitudinal bias of an appropriate strength is normally used to suppress the multidomain formation in the free layer of spin-valve sensors [55, 116, 176, 177, 178, 179]. Most of the longitudinal biasing techniques for spin-valve sensors stems from the earlier work on AMR sensors, and may be divided into two groups. The first group is based on exchange bias between a ferromagnet and an antiferromagnet [see Fig. 18(a)–(d)], and the second group is based on the magnetostatic interaction or exchange coupling between a ferromagnetic soft film and a permanent or hard magnet [see Fig. 18(e)–(h)].

4.2.1. Permanent Magnet-Based Bias

Contiguous Junction So far, the most widely studied bias scheme which is also employed in real disk drives is the contiguous (or abutted) junction using a permanent magnet (typically, a CoCrPt alloy), as shown schematically in Figure 18(e) [180–182]. The key to forming a proper bias in this scheme lies not only in the selection of a proper material with an appropriate thickness, but also in the control of the junction shape between the permanent magnet

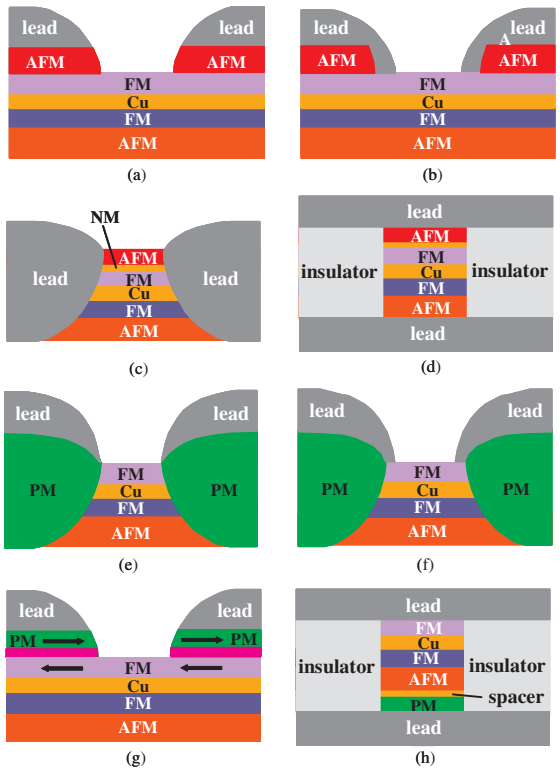


Figure 18. Schematics of different types of longitudinal bias schemes. (a) Patterned exchange. (b) Lead-overlaid patterned exchange. (c) Long-distance exchange. (d) Same as (c), but in CPP mode. (e) Contiguous junction with PM. (f) Lead-overlaid contiguous junction with PM. (g) Patterned synthetic ferrimagnet. (h) Parallel PM. The acronyms used are: PM—permanent magnet, FM—ferromagnet, AFM—antiferromagnet, NM—nonmagnetic layer. In (a)–(d), the top and bottom AFMs are two different types of materials.

and the active element of the sensor [183–188]. The overhanging of the permanent magnet on top of the free layer should be minimized; otherwise, it will cause hysteresis or kinks in the transfer curves, resulting in instability in the readback signal [189–191]. The ratio between the $M_r t$ of the permanent magnet and the $M_s t$ of the free layer or the soft layer should be optimized so as to obtain a good tradeoff between sensitivity and stability. The typical value is about 1.1–1.4 or even higher, depending on the shield-to-shield distance [123, 186, 192–194]. A small shield-to-shield distance requires a stronger permanent magnet. The microstructure of the permanent magnet also affects the performance of the sensor. The shape and cleanness of the junction are important for forming low-contact-resistance junctions, which is the key to suppressing the temperature rise of the reader [195]. To avoid oxidation and contamination at the junction interface, the permanent magnetic and contact are normally formed using a self-aligned lift-off process using a "mushroom" resist pattern formed on top of the sensor element [178, 179, 184]. The CoCrPt layer is typically grown on a Cr seed layer to have a higher coercivity, typically 1500–2500 Oe [196]. The uniform coverage of the seed layer on the junction edge is crucial to obtain a high-coercivity film along the entire junction. Poor seed-layer coverage may result in instability in the MR curve of the read sensor [197].

Lead-Overlaid Patterned Permanent Magnet The contiguous junction longitudinal bias is by far the most widely used bias scheme in real products. The drawback of this scheme, however, is that the bias field usually is not uniform across the longitudinal direction of the sensor. It is normally stronger at the two edges and weaker at the center. If the center portion is properly biased, then it is unavoidable that the edge regions will be overbiased, leading to the formation of so-called dead regions. These inactive regions will, in general, degrade the sensitivity of the senor. The influence of the dead region becomes more prominent when the read track narrows, in particular when the sensor width decreases to below 0.3 μm [198, 199]. An alternative scheme which can suppress the effect of the inactive region is the so-called lead overlaid structure in which the contact electrodes are extended over the abutted junction, and thus form a direct electrical contact with the inactive region of the sensor [Fig. 18(f)] [200]. Zhang et al. carried out a comparative study of magnetic noise in read heads with a contiguous junction and a lead overlaid design [201]. It was found that, at 0.25 μm magnetic read width and 4 mA bias current, the magnetic noise is twice as large as Johnson noise for the lead overlaid design, while it is comparable to Johnson noise for the contiguous junction design. The higher magnetic noise is attributed to a weaker longitudinal bias field with the lead overlaid design. This may affect the performance of read heads with a lead overlaid design at higher areal densities.

Patterned Synthetic Ferrimagnet Recently, the author's group proposed an alternative bias design in which patterned magnets are placed on top of the two edges of the free layer via an ultrathin Ru film such that the permanent magnets are antiferromagnetically coupled to the free layer [Fig. 18(g)] [202]. In conventional overlaid permanent magnet design or contiguous junction design with overlapping areas, the domain structure near the edges of the free layer is not so stable because the magnetic field generated by the permanent magnet inside the free layer always changes sign at the edges. In the new design, however, the antiferromagnetic coupling between the permanent magnet and the free layer forms a flux closure surrounding the permanent magnet, and thus improves the domain stability near the edges. Micromagnetic modeling has shown that, in addition to the domain stability, the sensitivity of the sensor with the new design is also higher than that of the contiguous junction design.

Permanent-Magnet-Based Design for CPP and MTJ Sensors The contiguous and lead overlaid designs are not suitable for CPP and MTJ sensors unless insulating permanent magnets are used to form the longitudinal bias or an additional insulating layer is added between the permanent magnet and the electrode at two edges. Redon et al. proposed a modified version of contiguous junction or patterned permanent design for an MTJ head through removing the portion of the pinned layer that is overlapped with the permanent magnet [203]. The results of micromagnetic modeling showed that the modified design is able to improve the performance of the MTJ head through suppressing the influence of the inactive region underneath the permanent magnet. Sin et al. proposed a design in which the longitudinal bias is formed through the magnetostatic coupling

from a permanent magnet placed at the other side of the pinned layer, as shown schematically in Figure 18(h) [204]. This design is suitable for CPP and MTJ sensors because there are no conducting materials at the two sides of the sensor element.

4.2.2. Antiferromagnet-Based Bias

Patterned Exchange As with the case of permanent magnet-based longitudinal bias design for spin valves, the exchange-biased designs also stemmed from the works on MR sensors [178, 179, 205, 206]. Among the exchange-domain stabilization designs, the most widely investigated is the patterned exchange design in which a patterned layer of antiferromagnetic material is fabricated on top of the free layer before the electrical lead is deposited [see Fig. 18(a)] [199, 207, 208, 209]. The advantage of this structure over the overlaid abutted junction structure is that the dead region is almost removed from the edge. In addition to this, the reduction in parasitic resistance associated with the junction improves the signal-to-noise ratio. However, in this case, one may need to use additional process steps to set the pinning direction of the patterned AFM layer because the exchange-bias direction of longitudinal bias for the free layer is 90° away from that of the traverse bias for the pinned layer (see Fig. 19). Another potential disadvantage of this design is the side reading due to the portions of the sensors underneath the patterned AFM layers. This might be suppressed by increasing the exchange-bias strength and the conductivity of the electrical contacts. It might also be reduced using the lead overlaid structure shown in Figure 18(b).

Long-Distance Exchange The most straightforward way to stabilize the domains of the sensing layer of MR and GMR heads is to add an AFM layer on top of the sensing layer. In fact, this has been employed in the early design of an MR head [177]. This design, however, lacks the flexibility in precise control of the bias field. It was often found that the exchange bias is too strong, which stiffens the magnetization of the sensing layer from rotation into the traverse direction, resulting in a low sensitivity. However, recently, there has been a revived interest in this design owing to the recent development of our understanding of the exchange bias and the ability to control the bias field through the insertion of an ultrathin nonmagnetic spacer at the AFM/FM interface [210–213]. Using Cu as the spacer, Nakashio et al. and Mao et al. successfully applied this design to a CIP spin-valve head and an MTJ head, respectively [Fig. 18(c) and (d)] [210, 211]. This type of design is particularly suitable for a CPP head.

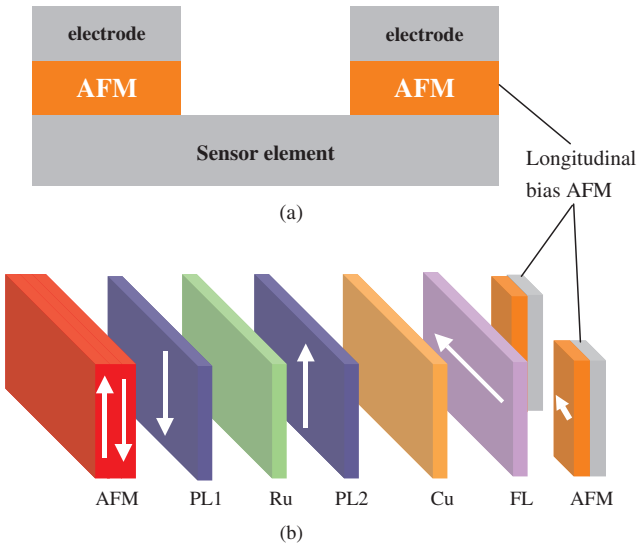


Figure 19. Cross-sectional view (a) and three-dimensional schematic structure (b) of spin valves with patterned-exchange longitudinal bias.

5. ELECTRONIC ASPECTS OF SPIN VALVES

Analogous to a transistor, the traverse and longitudinal biases discussed above are to set a proper working point for the spin valves. As is the case with transistors, once a proper working point is chosen, the performance of a spin valve is largely determined by its intrinsic properties, that is, the magnitude of the MR caused by the spin-dependent scattering at both the interfaces and the ferromagnetic layers. Although the spin valve exhibits a much higher MR as compared to the AMR sensor, the ever-increasing areal density requires continuous improvement of the performance of the spin valve. To this end, a variety of spin valves have been proposed and studied both theoretically and experimentally. These include, but are not limited to: (1) a multiple spin valve (e.g., dual spin valve), (2) a spin-filter spin valve, (3) a specular reflection spin valve, (4) a nanooxide-added spin valve, and (5) a nanooxide-added dual spin valve. In what follows, we explain briefly the motivation and underlying mechanism for each of these different types of spin valves illustrated schematically in Figure 14.

5.1. Spin Filter Spin Valve

The magnetoresistance of a CIP spin valve can be thought of as originating from the difference in the mean-free path (MFP) of the majority and minority electrons, at least phenomenologically. By making use of this phenomenon, Gurney et al. derived a novel method to measure the MFP of electrons in different materials such as Cu, AuCu, Co, NiFe, and Fe [214]. In this technique, a back layer is added to a spin valve with a thin free layer, and its effect on ΔG $(=G_P - G_{AP})$ is studied systematically at different back-layer thicknesses. Here, G_P and G_{AP} are the conductances of the spin valve in parallel and antiparallel configurations, respectively. It was found that, to a good approximation, ΔG can be written as

$$\Delta G = \Delta G_f + \Delta G_B[1 - \exp(-t_B/\lambda_+)] \qquad (24)$$

where ΔG_f is the conductance change due to the free layer, t_B is the thickness of the back layer excluding the inactive layer, if any, at the interface with the free layer, λ_+ is the MFP of majority electrons in the back layer, and ΔG_B is the conductance change due to the back layer when $t_B \gg \lambda_+$. The MFPs measured by this technique for Cu, Co, NiFe, and Fe are 19.2, 5.5, 4.6, and 1.5 nm, respectively. This result

implies that the addition of a back layer with an appropriate thickness can enhance the MR of a spin valve with an ultrathin free layer (so-called back-layer effect). The idea behind this is that the back layer helps to retain or even enhance the MFP of majority electrons, while it has little effect on the MFP of minority electrons. Kamiguchi et al. extended this concept to synthetic spin valves, and called them spin-filter spin valves [see Fig. 14(h)] [215]. The back layer is retermed a high-conductance layer (HCL). Note that the synthetic pinned layer is not shown in the schematic drawing in Figure 14(h) because we want to focus on the pure electronic effect. At a free layer thickness of 2–4 nm, the MR ratio can be enhanced by about 10% using the spin-filter design [216]. It is worth noting that, in addition to the enhancement of the MFP of majority electrons, the back layer or HCL also helps to better balance the current-induced fields at the free layer due to the redistribution of a certain portion of current to the HCL, leading to a lower asymmetry of the readout signal. For example, at a free layer thickness of 4 nm, Ueno et al. showed that the readout asymmetry can be reduced from 10% for a conventional spin valve to 3% for a spin-filter spin valve [216]. However, the improvement of the MR in a spin-filter design is only up to a certain extent because the HCL also increases the shunting current, which in turn will reduce the MR. Therefore, in most practical cases, the back layer is employed in combination with other techniques such as the nanooxide to improve the performance of advanced spin valves.

5.2. Dual Spin Valve

The most straightforward way to improve the MR ratio of a spin valve is to increase the number of free/pinned layer pairs. This has prompted the proposal and development of dual spin valves, in which the free layer is shared by two pinned layers at both sides (see Figs. 14(c) and 20) [217–223]. The MR of a typical dual spin valve can be 30–60% higher than that of a single spin valve, depending on the materials, device structure, and fabrication processes [21, 217, 218, 224]. The fabrication of a dual spin valve requires the use of AFM materials which are able to form exchange bias both at the top and bottom of the FM layer, and with a small critical thickness. Special cautions are needed in choosing the thickness of the two Cu spacers because two identical spacers may make it difficult to adjust the traverse bias field in the free layer using the sense current once the sensor fabrication is completed [223].

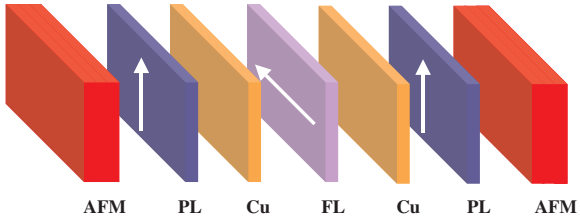


Figure 20. Schematic of a dual spin valve.

5.3. Specular Spin Valve

The dual spin valve consists of one free layer and two pinned layers. Can we increase the MR further by simply increasing the number of free and pinned layers? The answer is negative because different free layers may respond to the external field in a slightly different way, which may result in extra noises in the readout signal or degradation of the sensitivity. The only way to overcome this is to use antiferromagnetically coupled ferromagnetic/nonferromagnetic pairs. But as discussed earlier, this type of sensor is only applicable to the detection of a large magnetic field. One of the possible ways to increase the effective number of free layers without the need to add more layers physically is to increase the specular reflectivity of electrons at the outer surface of the free and pinned layer, respectively. As shown schematically in Figure 21, if the specular reflectivity at both surfaces is unity, the trilayer can be considered effectively as the repetition of the same unit in an infinite number of cycles. However, the difference between the equivalent structure and the real structure of the same type is that the former only has one free layer physically, and thus does not suffer from the drawbacks caused by the existence of a plural number of free layers.

5.3.1. Specular Spin Valve Using Insulating Antiferromagnet

The most straightforward way to implement this kind of structure is to use insulating AFM layers such as NiO and α–Fe_2O_3 to form an AFM/active layers/capping layer, AFM/active layers/AFM or AFM/active layers/NM/AFM sandwich structure (Fig. 14(j)). This has been verified experimentally to be very effective in increasing the MR ratio [219, 225–229]. The specular reflection enhancement effect was first described quantitatively by Swagten et al. in spin valves with structure NiO(50 nm)/FM1/NM1/FM2/NM2/NiO(10 nm), with FM1 = Co (2 nm), NM1 = Cu (2 nm), FM2 = Co (t), NM2 = Cu (1.2 nm) [225]. The second nonmagnetic layer, in this case, a 1.2 nm thick Cu, is used to isolate the exchange coupling between the top NiO and the second ferromagnetic layer (FM2), so that the top NiO functions only as a pure insulator. The decoupling effect of NM2 has been confirmed in the measured M–H curves, in which FM2 was found to switch at almost zero field. Using this structure, Swagten et al. obtained an MR ratio as high as 25% at 10 K and 15% at room temperature [225]. A semiclassical model has been invoked to interpret the experimental results, which suggests that the

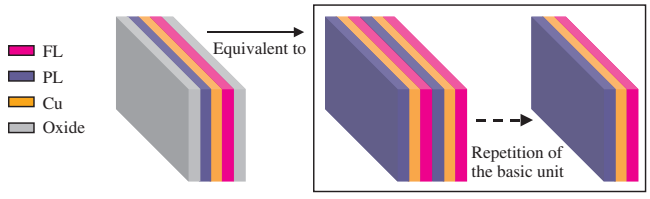


Figure 21. The number of the free layer (FL) and pinned layer (PL) can be increased effectively by increasing the specular reflectivity of both the top and bottom surfaces to unity.

enhancement of the MR is attributed to the enhancement of specular reflection at the metal/insulator interfaces. Also using the same NiO oxide, but in a dual spin valve with a structure NiO50/Co2.5/Cu1.8/Co4/Cu1.8/Co2.5/NiO 50 (thicknesses are in units of nanometers), Egelhoff et al. obtained an MR ratio of 21.5% at room temperature [219]. The large MR ratio was attributed to both the dual spin-valve structure and the enhanced specular scattering of electrons at the Co/NiO interfaces. Although NiO has a desirable electric property, its low exchange bias field and low blocking temperature make it an unfavorable choice for real read sensors. Hasegawa et al. carried out the first study on the possibility of using α–Fe_2O_3, another insulating AFM, to construct specular spin valves [226]. The MR ratio that has been obtained for a spin valve with structure glass/α–Fe_2O_3/$Ni_{80}Fe_{20}$6/Cu2.2/$Ni_{80}Fe_{20}$9/Cu4 (thicknesses are in units of nanometers) was only about 3.7%. Sugita et al. [227] and Kawawake et al. [228] grew α–$Fe_2O_3$50/Co2/Cu2/Co5/Cu0.4 (thicknesses are in units of nanometers) spin valves on a polished (110) plane of an α–Al_2O_3 single-crystalline substrate. The MR ratio obtained was as high as 18%. It was further increased to 27.8% for a dual spin valve with a structure α–Al_2O_3/α–$Fe_2O_3$50/Co2/Cu2/Co5/Cu2/Co2/α–$Fe_2O_3$50 (thicknesses are in units of nanometers). In either case, however, the unidirectional exchange-bias field is very small. The pinned layer is more or less a hard magnet instead of a soft layer pinned by an AFM. The coercivity of the bottom Co layer is more than 1200 Oe. The large MR ratio again was attributed to the enhanced specular reflection at α–Fe_2O_3/Co interfaces.

In addition to oxides, metals can also be used as the capping layer to increase the MR of spin valves. Although the mechanism of MR enhancement is not fully understood, it is believed that the potential step at both the free layer/capping layer interface and the capping layer/air interface may have helped to enhance the specular reflection of electrons. Compared to oxide capping layers, the metallic capping layer has certain limitations in thickness because too thick a capping layer will cause an increase in the shunting current, and thus leads to a decrease in the MR ratio [230, 231]. Although it has not been investigated in detail, this type of structure might be useful for CPP spin valves.

5.3.2. Specular Spin Valve Using Insulating Ferromagnet

Recently, Carey et al. [175] and Maat et al. [232] used insulating cobalt ferrite as pinning layers to improve the performance of spin valves. Spin valves with a high MR ratio (12.8%) and a large pinning field (about 1500 Oe) have been successfully fabricated. In this case, pinning is realized through an exchange-spring type of coupling between two ferromagnets [see Fig. 14(g)].

5.3.3. Nanooxide Added Spin Valve

As discussed above, the use of oxide AFM can greatly enhance the MR of spin valves. These oxides, however, are unsuitable for practical applications due to their low exchange bias field, low blocking temperature (for NiO), or large thickness required. This has led to the proposal of inserting nanooxides at the middle of the pinned layer of bottom spin valves. This, in combination with the top oxide capping layer, was found to greatly enhance the MR of spin valves. Since the pioneer work of Kamiguchi et al. [215], the nanooxide effect has been studied intensively to enhance both the magnetic and electronic properties of advanced spin valves [233–264]. The nanooxide, in general, is inserted at the middle of the pinned layer, and in the case of synthetic spin valves, it is normally inserted at the middle of the pinned layer that is nearer to the spacer. It is more effective when being used together with an insulating capping layer, leading to the so-called double specular spin valve [see Fig. 14(k)]. The nanooxide can be formed using different techniques such as natural oxidation, plasma oxidation, remote plasma oxidation or so-called atomic beam oxidation, ion beam oxidation, and ion-assisted oxidation. The plasma- and ion-based techniques were found to be able to form denser oxides than natural oxidation. Under optimum conditions, the oxides improve the MR without compromising the magnetic and thermal stability and the ESD robustness of spin valves. Read heads with sensitivity in the range of 6–10 mV/μm or higher have been demonstrated using nanooxide-added spin valves [257]. As the results obtained from different works are, to a certain extent, dependent on specific experimental conditions, the results are not always consistent with one another. Table 2 summarizes the representative works on specular spin valves reported so far in the last few years. Due to the large number of publications, the author cannot guarantee that the collection is complete. In what follows, we use our own results as examples to describe the main effect of nanooxide.

The Role of Oxide in Pinned Layers The roles of the bottom oxides are twofold. First, they reduce the surface roughness of the subsequently deposited layers, and second, they enhance the specularity of the pinned layer. The improvement of the surface roughness has been confirmed through both direct measurement of the surface roughness using AFM and measurement of the interlayer coupling field between the free and pinned layers [249]. In general, there are two major coupling fields between the free and pinned layers of different origin: (1) orange-peel coupling, and (2) RKKY coupling. The former is dominant in samples with rough interfaces, while the latter becomes more important when the interfaces become atomically flat. The orange-peel coupling decays exponentially with the thickness of the spacer, and for a sinusoidal roughness profile, it is given by [265]

$$H_{\text{int\_OP}} = \frac{\pi^2 A^2}{\sqrt{2} t_F \lambda} M_F \exp(-2\pi\sqrt{2} t_s/\lambda) \tag{25}$$

where A and λ are the amplitude and wavelength of the roughness profile, t_F and t_s are the thicknesses of the free and spacer layer, respectively, and M_F is the magnetization of the free layer. The original concept was extended to thin films with a structure of hard/spacer/soft by Kools et al. in 1999 [266]. Although Eq. (25) is modified slightly, the exponential dependence of the coupling field on the spacer thickness is still the same.

Table 2. Summary of the findings reported in nanooxide-related works.

Research group	Structure of spin valve (thicknesses are in nm)	GMR (%)	H_{in} (Oe)	H_{ex} (Oe)	Remarks	Ref.
1. NIST, Univ. Minnesota, Univ. California	NiO 50/Co 2.5/Cu 1.8/Co 4/ Cu 1.8/Co 2.5/NiO 50	21.5			First experimental indication of specular reflectivity enhancement	[219]
	NiO 50/Co 2.5/Cu 1.9/Co 4/ Cu 1.9/Co 2.5/NiO 50	24.8	Oscillating between 80 and −20 Oe		Systematic study of the role of background gas (CH_4, N_2, H_2O, C, H_2, CO, O_2);	[233]
	NiO 50/Co 2.5/Cu 2/Co 3/ CoO 0.4/TaO 0.4	19%			Oxygen was found to act as a surfactant, improving the properties of SV	
2. Eindhoven Univ. Tech., Philips	NiO 50/Co 2/Cu 2/Co (t)/ Cu 1.2/NiO 10	25 at 10 K 15 at R.T.			First quantitative study of specular reflection effect both experimentally and theoretically	[225]
3. Alps	glass/α–Fe_2O_3/$Ni_{80}Fe_{20}$ 6/ Cu 2.2/$Ni_{80}Fe_{20}$ 9/Cu 4	3.7%			First report on SV using/α–Fe_2O_3 as AFM	
4. Matsushita	α–Fe_2O_3 50/Co 2/Cu 2/Co 5/ Cu 0.4	18%			α–Fe_2O_3 was grown epitaxially on polished (110) plane of an α–Al_2O_3 single crystalline substrate	[227]
	α–Fe_2O_3 50/Co 2/Cu 2/ Co 5/Cu 2/Co 2/ α–Fe_2O_3 50	27.8%				
	Ta 3/PtMn 15/CoFe 2/Ru 0.7/ CoFe 1/NOL/CoFe 1.5/ Cu 2.2–2.9/CoFe 1/ NiFe 1–2/NOL/Ta 3	13		1092	Good thermal stability	[245]
	Ta/AF (PtMn or α–Fe_2O_3)/ CoFe/NOL/CoFe/Cu/ CoFe/Ta or Ta/AF (PtMn or α–Fe_2O_3)/ CoFe/NOL/CoFe/Cu/ CoFe/NOL/CoFe/Ta	12–17.3		414–723		[234]
5. Nagoya Univ., Matsushita	Ta 3/NOL/Co 2/Cu 2/Co 2/ α–Fe_2O_3 50/Al_2O_3	21.6			α–Fe_2O_3 as the AFM SV with NOL is more sensitive to temperature below 200 K	[235]
6. Toshiba	Ta 5/NiFe 2/IrMn 7/PL (with NOL)/Cu 2/CoFe 2/ CoFeO 0.4/TaO 0.4	16		400	First report on NOL effect inside the SV stack; Natural oxidation; Free layer Hc, 14Oe	[215]
	underlayer/IrMn 7/CoFe/ pin-NOL/CoFe 2/ Cu 2.2/CoFe 2/Cu 1/ free-NOL	17	7	400	ΔR_s: 2.5–3 Ω: Natural oxidation; Testing head for 10 Gbits/in^2; Promising for 100 Gbits/in^2	[243]
	Sub/Ta 3/$Ni_{62}Fe_{16}Cr_{22}$ 3/ AF (PtMn or IrMn)/ $Co_{90}Fe_{10}$ 1/$Co_{50}Fe_{50}$ 1 (for FeCo NOL), or $Co_{90}Fe_{10}$ 1 (for CoFe NOL)/ oxidation process/ $Co_{90}Fe_{10}$ 2/Cu 2.3/ $Co_{90}Fe_{10}$ 2/Cu 1/Ta 1	15.7	−2	500	NOL formed by ion-assisted oxidation is more stable than those formed by natural oxidation; Thick NOL reduces ferromagnetic coupling across NOL	[259]

continued

Table 2. Continued

Research group	Structure of spin valve (thicknesses are in nm)	GMR (%)	H_{in} (Oe)	H_{ex} (Oe)	Remarks	Ref.
7. Alps + Toshiba	Double specular (Ox–p + Ta–Ox) SAF bottom SV	17–18		1100–1500	Head level characterization including ESD	[258]
8. Fujitsu	Underlayer/$Pd_{32}Pt_{17}Mn_{51}$/ CoFeB Ru/CoFeB/Cu/ CoFeB/Cu/oxide	12.3	Oscillating between 10 and −18 Oe	>1000	Head for 56.1 Gbits/in^2	[260]
	Underlayer/$Pd_{32}Pt_{17}Mn_{51}$/ CoFeB Ru/CoFeB/ NOL/CoFeB/Cu/ CoFeB/Cu/oxide	15.1	Oscillating between 20 and −15 Oe		Head for 106.4 Gbits/in^2	
	Underlayer/PdPtMn/CoFe Ru/CoFe/NOL/Cu/ CoFe/NiFe/Cu/oxide	12.4	Oscillating between 20 and −15 Oe	>1000	Head for 106.4 Gbits/in^2	[241]
	Underlayer/PdPtMn/ CoFeB/Ru/CoFeB/ NOL/CoFeB/Cu/ CoFeB/Cu/oxide	15.1	Oscillating between 20 and −20 Oe		Head for 56.1 Gbits/in^2	[242]
	Underlayer/PdPtMn/ CoFeB/Ru/CoFeB/Cu/ CoFeB/Cu/oxide	12.3	+2	~2000	$t_{Cu} = 20$	
	Underlayer/PdPtMn/ CoFeB Ru/CoFeB/Cu(t)/ CoFeB(1.5)/Cu/oxide	8.87	Oscillating (~−17 Oe when $t_{Cu} = 2$–3 nm)	938	Capping layer effect: Ta capping: MR ratio, 7.55%, interlayer coupling field, 5.1Oe; Oxide capping: MR ratio, 8.87%, interlayer coupling field, −16.6%	[236]
9. Philips	Ta 3.5/NiFe 2/IrMn 1/ CoFe 3/NOL/CoFe (t)/ Cu 2.5/CoFe 4/NOL	14		~800	Natural oxidation; Al_2O_3, a good top NOL	[248]
	Ta 3.5/NiFe 2/IrMn 8/CoFe 3/ Cu 2.5/CoFe 4/NOL	9.5			CoFe is preferred over Co because the former allows for the formation of iron oxide	[239]
	Ta 3.5/NiFe 2/IrMn 8/ CoFe 3/NOL/CoFe 3 Cu 2.5/CoFe 4/NOL	13				
10. INESC	Ta 6.7/NiFe 4.2/IrMn 9/ CoFe 1.4/NOL/CoFe 1.5/Cu 2.2/CoFe 4/ NOL/Ta	12.5	3		Temperature-dependent study; Anomalous bumps seen in MR–T curves of NOL SV	[251]
	Ta 7/NiFe 5/IrMn 9/ CoFe 1.4/NOL/CoFe 2.5/ Cu 2/CoFe 2.6/ NOL/Ta 1–2	13.6	<4	~180	Remote plasma oxidation; RBS result indicates a thickness of 1.5 nm for NOL	[240]
11. TDK	Sub/Ta 5/NiFe 2/NOL/ NiFe 1/CoFe 2/ Cu 2.5/CoFe 2/NOL/ CoFe/PtMn or RuRhMn/Ta 5	11.8 (RuRh Mn)			Better thermal stability of SV with NOLs	[237]
		9.4 (PtMn)				
12. Sony	Glass/Ta 3–5/NiFe 12/PtMn 10–20/CoFe 1.1–1.5/ Ru 0.8/NOL/CoFe 2–2.2/ Cu 1.9–3.2 CoFe 1.5–2.5/ Cu 1/Ta (Ta–O) 0.2–5	14.9			Ru/Ox Good thermal stability	[261]

continued

Table 2. Continued

Research group	Structure of spin valve (thicknesses are in nm)	GMR (%)	H_{in} (Oe)	H_{ex} (Oe)	Remarks	Ref.
13. IBM	Al_2O_3 3/NiCrFe 3/NiFe 1/ PtMn 20/CoFe 2/ Ru 0.8/CoFe 2.2/ Cu 2.2/CoFe 0.9/ NiFe 2.7/Cu 8/ Al_2O_3 1/Ta 6	13.8	Oscillating between 25 and −12 Oe	~3000	>20 Gbits/in^2: Al_2O_3 + NiCrFe as the seed layer; Strong {111} texture; Good thermal stability; Head level test (good linearity & small asymmetry (−1.5%), 6.64 mV/μm, 20 Gbits/in^2)	[244]
14. IBM, San Jose State Univ.	Ta 3–4/Cu 1/IrMn 8.5/CoFe (d)/NOL/CoFe (3.3-d)/ Cu 2.4/CoFe 1.6/ NiFe 2.2/Ta 3.5	10.4	5.1	460	Natural oxidation; MR ratio highest when d ~ 1.2 nm	[262]
	Ta 3–4/Cu 1/IrMn 7/CoFe (d)/NOL/CoFe (3-d)/ Cu 2.4/CoFe 3/Cu 2.7/ CoFe 3/IrMn 7/Ta 3.5	20.5				
15. ReadRite	Ta 1.5–3/NiFeCr or NiFe 2–5/ PtMn 10–20/CoFe 1.5–2/ Ru 0.8/CoFe 1/NOL/ CoFe 1.5–2/Cu 1.5–2.4/ (CoFe/NiFe or CoFe) 2–3/ Ta/TaO 1–5	16.1 (plasma NOL), 15 (natural NOL), 12.5 (non-NOL)	Weak oscillation (>0) 5.8 at t_{Cu} = 2.2 nm	1000	>50 Gbits/in^2: Plasma oxidation forms denser and more uniform NOL; Optimum Ta capping layer thickness ~1 nm; No degradation in thermal stability; Head level test (6.5–10 mV/μm, asymmetry, <2% for positive sense current, good for 50 Gbits/in^2)	[257]
16. LLNL ReadRite	FM1/Cu/FM2 FM1 = NiFeCo, NiFeCo/ CoFe, CoFe, FM2 = NiCoFe, CoFe/ NiFe, CoFe/NiFe/CoFe	9.3 (non-NOL), 12.9 (NOL at one side), 13.9 (NOL at both side)			Nonexchange-biased sandwich	[238]
17. Veeco LLNL	Underlayer/NOL/NiFe 2.5/ CoFe 1/Cu (t)/CoFe 3/ IrMn 7/Ta 2	~12.5	Oscillation (>0)	~330–430	Systematic study of different oxides; MR increases from 26.5 to 36.7% in the order of AlO_x, TaO_x, CuO_x, NbO_x, CrO_x, $NiFeCrO_x$, $NiFeO_x$, $CoFeO_x$	[256]
18. Veeco, ReadRite, LLNL, Toshiba	SiO_2/Ta 5/NiFe 2/$Ir_{20}Mn_{80}$ 7/ CoFe 2/NOL/CoFe 2/ Cu 2/CoFe 2/Cu 1/Ta 1 Alternative structures (replacing $Ir_{20}Mn_{80}$ by $Pt_{50}Mn_{50}$ 15, or $Pt_{50}Mn_{50}$ 15/$Co_{10}Co_{90}$ 2/Ru 0.8, or $Ir_{20}Mn_{80}$ 7/$Co_{10}Co_{90}$ 2/Ru 0.8	15.5		~480	PVD vs. IBD; PtMn and IrMn; Simple and synthetic; Atomic beam oxidation gives the highest FM coupling across the NOL	[246]

continued

Table 2. Continued

Research group	Structure of spin valve (thicknesses are in nm)	GMR (%)	H_{in} (Oe)	H_{ex} (Oe)	Remarks	Ref.
19. Veeco + LLNL + Toshiba	Top SV: Ta 2/NOL/ NiFe 1/CoFe 2/ Cu 2.4/CoFe 2/ IrMn 6/Ta 2	9.8	Positive oscillation (5–30)	574	Systematic study of atomic beam oxidation (ABO) or remote plasma oxidation and ion beam oxidation (IBO); ABO gives largest MR ratio and exchange-bias field for bottom spin valves	[245]
	Bottom SV: Ta 5/NiFe 2/ IrMn 7/CoFe (2-t)/ NOL/CoFe 2/Cu 2/ CoFe 2/Cu 1/Ta 1	12.2 (ABO), 11.9 (ABO)		Increase from 300 to 350		
	Dual SV: Ta 2/NiFe 2/ IrMn 6/CoFe (1.5-t)/ NOL/CoFe 2/Cu 2/ CoFe 2.5/Cu 2/CoFe (2-t)/NOL/CoFe 1.5/ IrMn 6/Ta 2	18.5 (ABO)		~300		
20. Univ. Utah + LLNL	Ta 2/NOL/NiFe 2.5/CoFe 1/ Cu (t)/CoFe 3/IrMn 7/ Ta 1.5	12.25	29	~400	Natural oxidation; Cu wedge SV; MR decreases monotonically with Cu thickness	[263]
21. Seoul Nat. Univ. + KIST	Sub/Ta 5/$Ni_{81}Fe_{19}$ 2/ $Fe_{50}Mn_{50}$ 8/$Co_{90}Fe_{10}$ 1.5/ NOL/$Co_{90}Fe_{10}$ 1.5/ Cu 2.6/$Co_{90}Fe_{10}$ 2.5/ Cu 2.6/$Co_{90}Fe_{10}$ 1.5/ NOL/$Co_{90}Fe_{10}$ 1.5/ $Fe_{50}Mn_{50}$ 8/Ta 5	15.9		~250		[252]
	Sub/Ta 5/$Ni_{81}Fe_{19}$ 2/ $Fe_{50}Mn_{50}$ 8/$Co_{90}Fe_{10}$ 2/NOL/$Co_{90}Fe_{10}$ 2/Cu 2.6/$Co_{90}Fe_{10}$ 1.5/$Ni_{81}Fe_{19}$ 4.5/Ta 5	10.1			NOL may also act as a diffusion barrier for Mn	[253]
22. Nal. Tsing Hua Univ.	Top SV: Ta 8/Co 5/Cu 2.1/ Co 4/FeMn 10/Ta 5 Bottom SV: Ta 8/NiFe 5/ FeMn 10/Co 5/Cu 2.1/ NiFe 10/Ta 5	7.8			Bottom SV: NOL at FeMn/Co and NiFe/Ta interfaces increases and NOL at Co/Cu or Cu/NiFe interfaces decreases the MR ratio; Top SV: NOL at Ta/Co interfaces increases MR ratio	[264]
23. Inst. of Phys., CAS UST, Beijing Univ. Plymouth, Nordiko	Ta 3.5/NiFe 2/IrMn/6/ CoFe 1.5/NOL/CoFe 2/ Cu 2.2/CoFe (t)/NOL/Ta	15		~380	Bottom NOL is a mixture of metal and oxide	[255]
	Ta/NiFe/IrMn/CoFe NOL/ CoFe/Cu/CoFe/Cu/ NOL/Ta	15	~5	~400	Spin filter SV with NOL (already employed by other groups)	[254]
24. DSI/NUS	Top SV: Ta 3/NOL/NiFe 2/ CoFe 1.5/Cu 2.2/CoFe 3/ IrMn 8/Ta 2	10.8	5–15	225–450	MR increases in the order of $CoFeO_x$, TaO_x, CuO_x, AlO_x	[250]
	Bottom SV: Ta 3/NiFe 2/ IrMn 6/CoFe 2.5/Cu 2.2/ CoFe 1.5/NiFe 3/Cu 2/ NOL/Ta 1	8.8	5–24	520–555	MR increases in the order of TaO_x, AlO_x. $NiFeO_x$, $CoFeO_x$, CuOx	

continued

Table 2. Continued

Research group	Structure of spin valve (thicknesses are in nm)	GMR (%)	H_{in} (Oe)	H_{ex} (Oe)	Remarks	Ref.
	Dual SV: Ta 3/NiFe 2/ IrMn 6/CoFe 1.5/NOL/ CoFe 1.5/Cu 2/CoFe 1/ NiFe 1/CoFe 1/Cu 2/ CoFe 2/NOL/CoFe 1/ IrMn 6/Ta 2	21.8		~280	Best bottom oxide: $CoFeO_x$ Best capping oxide: AlO_x	
	Ta 3/NiFe 2/IrMn 6/CoFe 1.5/ NOL/CoFe 1.5/Cu 2.2/ CoFe 1/NiFe 2/CoFe/NOL/ Cu 1/Ta 1	15.3	Oscillating between 0 and 50	~350	Both specular reflection enhancement and interface smoothening effects have been confirmed, but interface layer coupling is still ferromagnetic	[249]

On the other hand, as discussed before, the strength of RKKY coupling oscillates with spacer thickness:

$$H_{\text{int\_RKKY}} \propto \frac{1}{t_F}\frac{1}{t_s^2}\sin\left(\frac{2\pi t_s}{\Lambda}+\phi\right) \tag{26}$$

where Λ is the oscillation period and ϕ is the phase factor. Except for a few reports [244, 257], the oscillation can hardly be seen in normal spin-valve samples without nanooxides. However, it can be observed easily in samples with nanooxides. The suppression of orange-peel coupling has also been achieved in Cu/Co superlattices through the introduction of a proper amount of oxygen during the growth of the entire film stack [233, 267].

To confirm both effects of nanooxide, spin valves with the structure Ta3/NiFe2/IrMn6/CoFe3/Cu(t)/CoFe1/NiFe2/CoFe1/Cu1/Ta1 [t denotes the thickness of the spacer layer; thicknesses are in nanometers, hereafter, we call them conventional SVs (CSV)] and SVs with the structure Ta3/NiFe2/IrMn6/CoFe1.5/NOL/CoFe1.5/Cu(t)/CoFe1/NiFe2/CoFe1/OX/Cu1/Ta1 [hereafter, we call them nanooxide-added SVs (NOL-SV)] were deposited on Si(100) substrates coated with a 1 μm thick thermally oxidized SiO_2 layer by using an ultrahigh vacuum sputtering system operating at a base pressure of 5×10^{-10} torr. Here, NOL denotes the natural oxidation of the CoFe layer. Oxidation was conducted in a separate chamber, by exposing the fresh CoFe surface to pure oxygen atmosphere. The exposure time was 1 min, but the pressure was controlled from 1×10^{-4} to 1.6 Pa. The specimens were magnetically annealed at 275 °C for 1 h under magnetic fields of 1 T in a commercial magnetic vacuum annealing oven under a pressure of 1.0×10^{-6} torr. The MR was measured by using a linear four-point probe method. The sheet resistance was calibrated by measuring the MR on a microstructural bridge with 400 μm length and 100 μm width.

Figure 22(a) shows the dependence of the MR ratio and interlayer coupling field on the oxygen exposure dose used to oxidize the CoFe layer. It is found that H_{int} decreases almost monotonically with the oxygen exposure dose, decreasing from 31 Oe in a CSV to 4 Oe in an NOL–SV. On the other hand, the MR ratio increases from 9 to 14% with an increasing O_2 exposure dose from 10^{-8} to 0.1 min Pa, which saturates when the exposure dose is over 0.1 min Pa. Shown in Figure 22(b) are the sheet resistance R_{sq} and its change ΔR_{sq} as a function of O_2 exposure dose. R_{sq} decreases initially, reaches a minimum at an oxygen exposure dose of about 10^{-3} min Pa, and increases when the oxygen dose increases. The initial decrease of the sheet resistance might be attributed to the enhancement of specular scattering of electrons at the interface between the metal and NOL.

Figure 23 shows the MR (a), interlayer coupling field (b), sheet resistance (c), and change of sheet resistance (d) as a function of Cu thickness. The enhancement of surface smoothness can be seen from the oscillation of the interlayer coupling field with the Cu layer thickness in NOL–SVs, as shown in Figure 23(b) (solid circle). Also shown in the figure are the experimental results for conventional spin valves (open circle) and the fitting curve according to

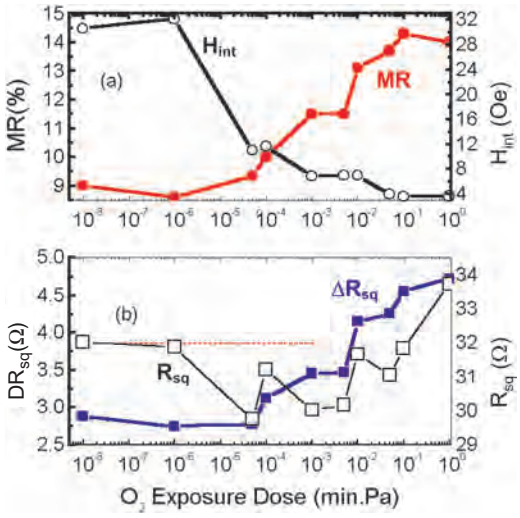


Figure 22. (a) MR ratio (solid circle) and H_{int} (open circles) as a function of O_2 exposure dose. (b) Sheet resistance (open squares) and change of sheet resistance ΔR_{sq} (solid squares) as a function of O_2 exposure dose for a series of SVs with the structure Ta3/NiFe2/IrMn6/CoFe1.5/NOL/CoFe1.5/Cu2/CoFe1/NiFe1/CoFe1.2/NOL/Cu1/Ta1 with a deferent O_2 exposure dose. The thicknesses are in nanometers.

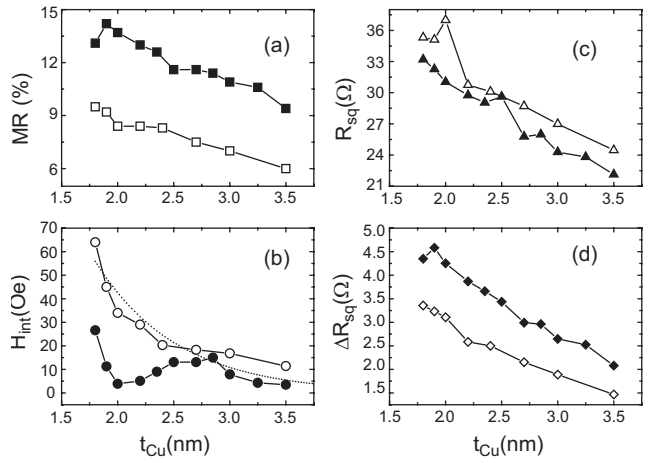


Figure 23. Spacer (Cu) thickness dependence of MR (a), H_{int} (b), sheet resistance R_{sq} (c), and sheet resistance change (d) for two batches of SVs with the structure Ta3/NiFe2/IrMn6/CoFe3/Cu(t)/CoFe1/NiFe2/CoFe1.2/Cu1/Ta1 (open symbols) and Ta3/NiFe2/IrMn 6/CoFe1.5/NOL /CoFe1.5/Cu(t) /CoFe1/NiFe2/CoFe1.2/NOL/Cu1/Ta1 (solid symbols). The dotted line in (b) was fitted by using Néel's orange-peel coupling equation. The thicknesses are in nanometers.

Eq. (25) (dotted line). Although the interlayer coupling is still ferromagnetic in the entire range of the Cu thickness that has been investigated, the clear oscillation in NOL–SVs and the monotonic decrease with Cu thickness in CSVs clearly demonstrate the effectiveness of NOL in improving the surface roughness. Figure 23(c) and (d) shows the dependences of the sheet resistance and its changes on the Cu thickness, respectively. It is interesting to note that the nanooxide-added spin valve has a smaller R_{sq}, but a larger ΔR_{sq} than the conventional spin valve in the entire thickness range for Cu, leading to a larger MR ratio in the whole range, as shown in Figure 23(a). It is also interesting to note that the maximum of MR corresponds to the minimum of the interlayer coupling field. This can be understood readily because, at this thickness range, the free and pinned layers are antiferromagnetically coupled if one subtracts the orange-peel coupling field from the total coupling field. Very recently, we have been able to achieve antiferromagnetic coupling routinely in spin valves with oxide capping layers.

Effects of Different Types of Oxides There are many different types of oxides that can be used to improve the performance of the spin valves. The selection of oxides that have been studied so far is more or less based on the availability of the elements that are used to form the oxides in individual deposition systems. In general, one has more choices for the capping layer than that for the NOLs inserted in the middle of the spin valves because the latter will affect the magnetic properties of the spin valves. For example, if it is inserted in the middle of the pinned layer, it cannot be too thick; otherwise, it will separate the pinned layer into two sublayers, and result in a substantial decrease of the exchange bias. Although it can be made thicker if it is inserted at the AFM/seed layer interface, it cannot be too thick; otherwise, it will affect the crystalline texture of the layers that are subsequently deposited on top of the oxide. This eventually will lead to a decrease of exchange bias. Bearing all of these in mind, the author's group has carried out a systematic study on: (1) how different types of oxides affect the performance of the spin valve, and (2) which position is more desirable in terms of performance and manufacturability [250].

To investigate the effect of different types of oxides, we fabricated simple top and bottom spin valves with oxides being added at different locations, as shown schematically in Figure 24. The oxides that we have investigated include $CoFeO_x$, TaO_x, AlO_x, $NiFeO_x$, and CuO_x. The effects of these oxides on the interlayer-coupling field H_{in}, exchange bias H_{ex}, and MR ratio are shown in Figure 25. The main results can be summarized as follows. First, the exchange bias of the top spin valve is much smaller than that of the bottom spin valve. This is caused by both the property of the AFM (in this case, IrMn) itself and the oxide-induced changes in the texture of the underlying layers. The latter is especially severe in the case of TaO_x and AlO_x. $CoFeO_x$ and $NiFeO_x$ are better because they were formed from the oxidation of CoFe and NiFe, which already had a well-established texture before the oxidation took place. So does Cu because it is usually used as the texture control layer for NiFe and CoFe. Second, the interlayer-coupling field of the top spin valve is, in general, smaller than that of the bottom spin valve, except for the case using AlO_x. This again suggests that the nanooxide has the surface smoothening effect, as discussed earlier. The improvement is particularly remarkable for $CoFeO_x$. Finally, we can see that the MR ratio of the top spin valve, in general, is larger than that of the bottom spin valve. Comparing these results, we can see that $CoFeO_x$ is, so far, the most desirable bottom oxide, while AlO_x seems to be a better choice for the top oxide. For other oxides to be better than $CoFeO_x$, one must find a material that has a similar crystalline structure and texture as that of the CoFe layer in spin valves.

Effect of the Position of Oxide In addition to the types of oxides, we have also carried out a systematic study on the position of oxide, that is, which position is more suitable for the bottom oxide. Figure 26 shows schematically the different positions that have been investigated. These are: (1) inside the AFM layer, (2) at the AFM/first pinned layer interface, (3) at the Ru/second pinned layer interface, (4) inside the second pinned layer, (5) at the second pinned layer/Cu interface, (6) inside the spacer, (7) at the Cu/free layer interface, and (8) at the top surface. In what follows, we describe the results for cases (4)–(6).

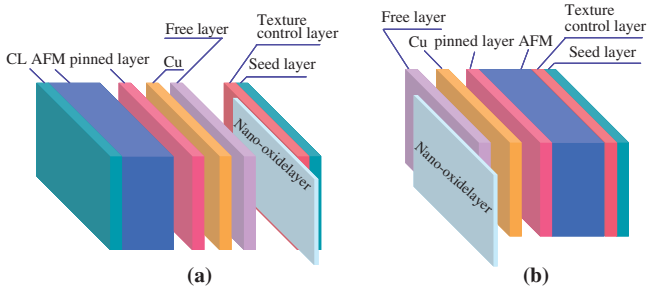


Figure 24. Schematic of a top spin valve with a nanooxide inserted at the free and texture control layer interface (a), and a bottom spin valve with a nanooxide capping layer (b).

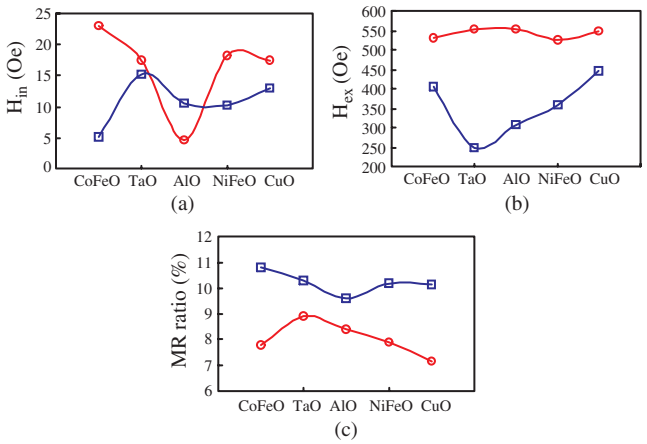


Figure 25. Effect of different nanooxides on the interlayer coupling field (a), exchange bias (b), and MR ratio (c) of bottom (open circles) and top (open squares) spin valves whose structures are shown in Figure 24.

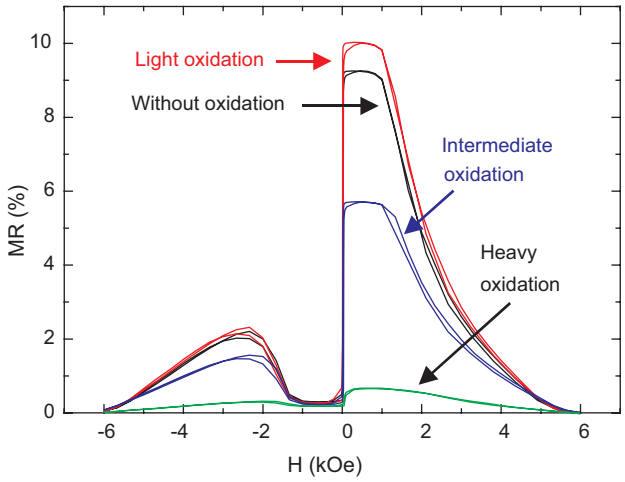


Figure 27. MR–H curves of spin valves oxidized under different oxygen doses at the surface of the second pinned layer.

Figure 27 shows the MR–H curves of case (5), in which the spin valve has a structure Ta3/NiF2/IrMn8/CoFe2.5/Ru0.8/CoFe2/NOL/Cu2.2/CoFe2/Cu1/Ta1.5 (thicknesses are in nanometers). Here, NOL stands for the nanooxide layer. The NOL was formed through natural oxidation of the second pinned layer in a separate chamber, with the thickness being controlled by the oxygen dose. After the oxidation was completed, the sample was brought back again to the growth chamber for the deposition of the rest of the layers. All of the deposition and oxidation processes were accomplished in a multiple-chamber ultrahigh-vacuum system without breaking the vacuum. As can be seen from the figure, the MR ratio initially increases under light-oxidation conditions, beginning to decrease when the oxygen exposure dose increases, and eventually vanishing out as the oxidation process continues. Figure 28 shows the detailed dependence of the MR ratio and the interlayer coupling on the oxygen dose. It is worth noting that the interlayer-coupling field decreases and the MR ratio increases in the light-oxidation regime by increasing the oxygen dose. However, the former increases sharply when the latter starts to decrease as the oxygen exposure dose is further increased. The behavior of the MR ratio is easily understood, but it is not as straightforward to explain the sharp increase of the interlayer-coupling field. As the interlayer-coupling field normally decreases by decreasing the pinned layer thickness [266], it is plausible to postulate that the sharp increase of H_{in} has something to do with the decrease of the spacer thickness because the surface roughness should not change so drastically. Figure 29 shows the X-ray diffraction spectra of spin valves under different oxidation conditions. Curves (a)–(d) correspond to cases (1)–(4) shown in Figure 28. The peak at an angle of about 41.2° is from the IrMn layer, while that at about 43.7° is from the rest of the layers. It is worth noting that the higher degree peak splits into two new peaks in case (d): one at an angle of about 42.9°, and the other at about 44°. Although the details are not well understood at present, it seems that Cu has been oxides partially to form copper oxide or a mixture of copper oxide and CoFe oxide. The sharpness of the interfaces has been investigated using low-angle X-ray reflectometry. As shown in Figure 30, the four curves corresponding to the four different cases shown in Figure 29 can hardly be differentiated from one another. This implies that the change in surface roughness is unlikely responsible for the sharp increase of the interlayer-coupling field shown in Figure 28. The results demonstrate that it is difficult to form a controllable oxide at the interface between two materials if both can be readily oxidized. Figure 31 shows schematically the possible oxidation processes. Under heavy oxidation conditions, the loosely absorbed oxygen atoms on the $CoFeO_x$

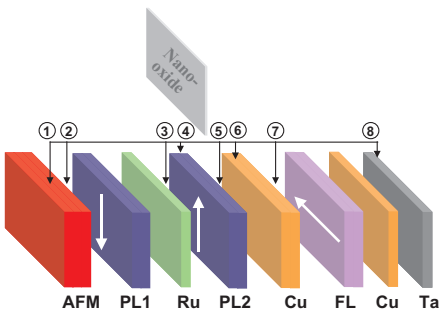


Figure 26. Schematic of the nanooxide-added spin valves with the oxide inserted at different locations. Here, PL1 and PL2 refer to the first and second pinned layers, respectively, and FL stands for the free layer.

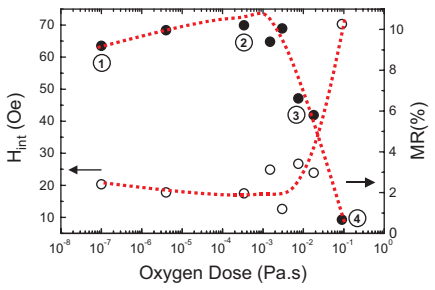


Figure 28. Dependence of the interlayer coupling field and MR ratio of spin valves on the oxygen dose that was used to oxidize the second pinned layer.

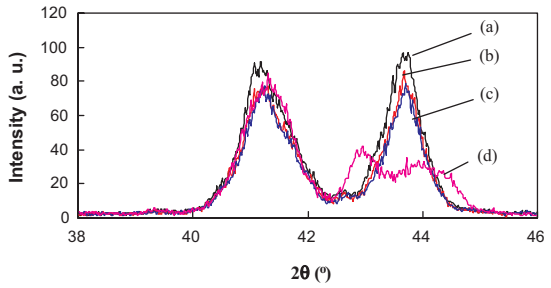


Figure 29. High-angle X-ray diffraction spectra of spin valves oxidized under different conditions. Cases (a)–(d) correspond to cases (1)–(4), respectively, in Figure 28.

surface may result in a partial oxidation of the subsequently deposited Cu layer. This may explain the sharp increase of the interlayer coupling field when the oxygen exposure dose exceeds a critical value.

From the above discussion, it is clear that adding NOL at the second pinned layer/Cu interface might not be a viable approach due to its narrow process window. The next position that has been widely studied since the first report of the nanooxide spin valve is in the middle of the pinned layer. Figure 32 shows the dependence of the interlayer-coupling field and the MR ratio on the oxygen exposure dose for which the spin valve has the structure Ta3/NiF2/IrMn8/CoFe2.5/Ru0.8/CoFe1.5/NOL/CoFe1.5/Cu2.2/CoFe2/Cu1/Ta1.5 (thicknesses are in units of nanometers). In this case, the oxide was formed in the middle of the second pinned layer. As can be seen from the figure, the interlayer-coupling field is almost constant at the beginning, decreases, and then increases again when the oxygen exposure dose increases. The MR ratio, however, increases first, then decreases, and increases again. Although it is difficult to understand the behavior of the MR, one thing that is clear is that the process window of adding oxide at this position is much wider than that of the previous case; thus, it is more desirable from the point of view of manufacturability.

The last position to be discussed here is the middle of the spacer. This position has not been discussed before, and thus it is worth investigating how it affects the performance of the spin valve. One tends to think that this will affect

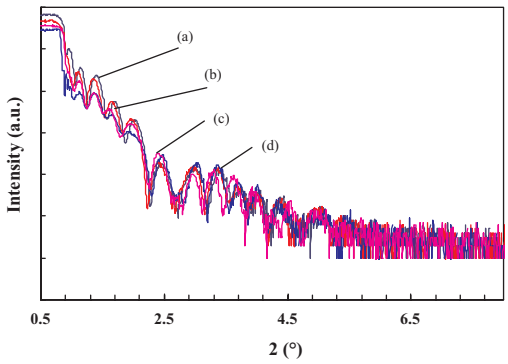


Figure 30. Low-angle X-ray diffraction spectra of spin valves oxidized under different conditions. Cases (a)–(d) correspond to cases (1)–(4), respectively, in Figure 28.

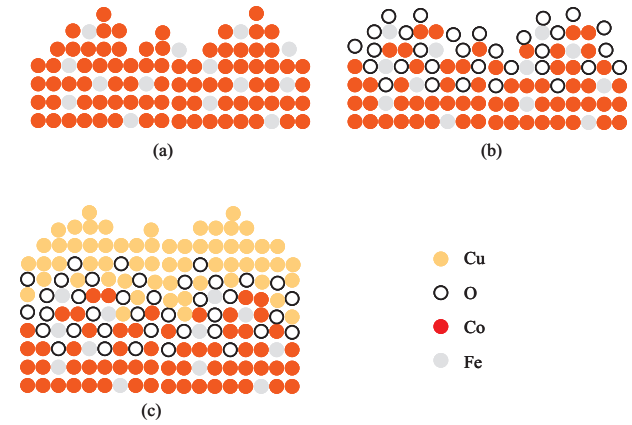


Figure 31. Schematic illustration of the oxidation process of CoFe (b), and the subsequent deposition and partial oxidation of Cu (c). (a) shows the CoFe surface before oxidation.

the performance adversely because it may reduce the mean-free path of electrons. However, as shown in Figure 33, the MR ratio indeed increases with the oxygen exposure dose that is used to oxidize the Cu layer in a spin valve with a structure Ta3/NiF2/IrMn8/CoFe2.5/Ru0.8/CoFe2/Cu1.2/NOL/Cu1.2/CoFe2/Cu1/Ta1.5 (thicknesses are in units of nanometers). This might be attributed to the enhancement of scattering probability at the free layer/Cu/pinned layer interfaces. In this case, the MR will increase provided that the scattering inside the spacer is spin conserved. Further study is required to understand the underlying mechanism.

As we mentioned above, the study of the nanooxide effect is still at an early stage. A theoretical framework has yet to be formulated to guide the selection of oxides and their positions in the spin valves. Further experimental investigations are required to study the structural properties of the nanooxide-added spin valves. Before ending this section, we show in Figure 34 the MR–H curve of a dual spin valve with a structure Ta3/NiFe2/IrMn5/CoFe1/NOL/CoFe2.3/Cu2.15/CoFe2/Cu2.15/CoFe2.3/NOL/CoFe1.0/IrMn5/Ta1. After the optimization of structure and processes, we have successfully obtained an MR ratio of 27.2%, which we believe is the highest value ever reported for all-metal spin valves. Although the exchange-bias field is too low for practical

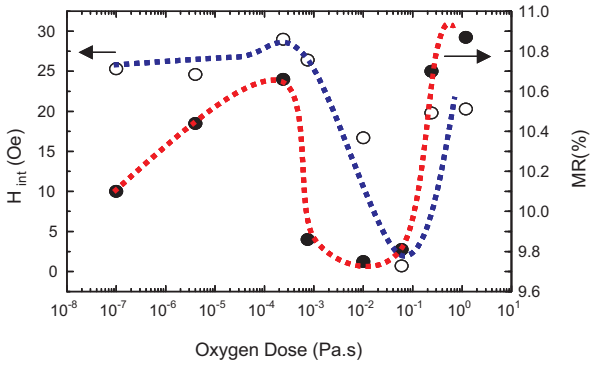


Figure 32. Dependence of the interlayer coupling field and the MR ratio of the spin valves on the oxygen dose that was used to form the oxide at the middle of the second pinned layer.

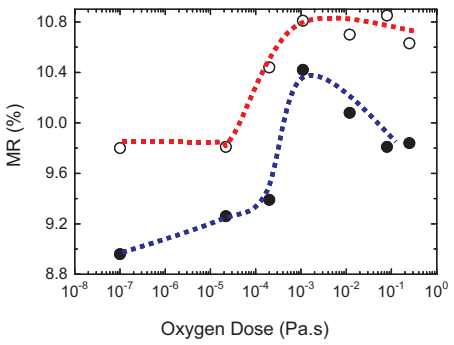


Figure 33. Dependence of the MR ratio of the spin valves on the oxygen dose that was used to form the oxide in the middle of the spacer. The upper curve was obtained from the major loop, while the lower curve was obtained from the minor loop. Both show a similar dependence on the oxygen dose.

applications, this might be solved by replacing IrMn with other AFM materials such as PtPdMn and NiMn.

5.4. CPP Spin Valves

5.4.1. Why CPP?

So far, most of the discussion has been concentrated on CIP spin valves. Although many of the fundamental concepts for CIP spin valves are directly applicable to CPP spin valves, there are some differences in how their performance can be improved. In this section, we first explain why we still need CPP spin valves for data storage, followed by a discussion on possible ways for improving the performance of CPP spin valves.

The areal density of a hard disk drive is the product of the linear density and track density. The former is determined by many component and system parameters, while the latter is largely determined by the capability of lithography techniques. In what follows, we first focus on linear density, which is defined as

$$\text{linear density} = \frac{1}{\text{bit length}} = \frac{\text{channel density}}{PW_{50}} \tag{27}$$

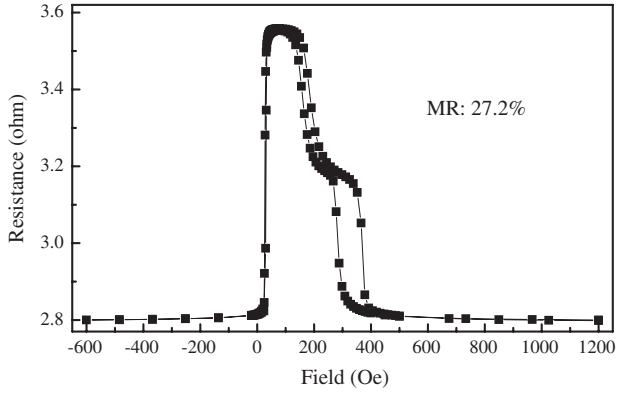


Figure 34. MR–H of a double specular dual spin valve with a structure Ta3/NiFe2/IrMn5/CoFe1/NOL/CoFe2.3/Cu2.15/CoFe2/Cu2.15/CoFe2.3/ NOL/CoFe1.0/IrMn5/Ta1. Here, the thicknesses are given in nanometers.

where PW_{50} is the full-width-at-half-maximum of the readout pulse from an isolated transition. The channel density is a parameter that reflects the efficiency of the channel coding scheme (~2.5–3). The PW_{50} can be estimated using the following equation:

$$PW_{50} = \left[2(g + t/2)^2 + 4(d_{\text{eff}} + a)^2\right]^{1/2} \tag{28}$$

where

$$d_{\text{eff}} = \sqrt{d(d+\delta)} \quad \text{and} \quad a = k\left(d_{\text{eff}}\frac{M_r t}{Hc}\right)^{1/2} \tag{29}$$

Here, d is magnetic spacing, g is the gap length between the sensor and the two shields, t is the thickness of the sensor, δ is the thickness of the media, and a is the transition width. M_r and H_c are the remnant and coercivity of the media, respectively. k is a parameter with a value ranging from 1 to 2. The above equation suggests that there are many possible ways to increase the linear density, although the reality is that none of them is an easy task. Although one can always try to push all of the parameters to their individual upper or lower limits, it would make more sense if one tries to determine which parameter is of more relative importance first. To this end, we have tried to find the rate of change of PW_{50} with respect to different parameters by assuming that we have already achieved 100 Gbits/in^2. It turned out that

$$\frac{\partial PW_{50}}{\partial d} = 2.6\frac{\partial PW_{50}}{\partial(M_r\delta/H_c)} \quad \text{and}$$
$$\frac{\partial PW_{50}}{\partial d} = 2.1\frac{\partial PW_{50}}{\partial(g+t/2)} \tag{30}$$

where the following parameters have been used for the calculation: $d = 12$ nm, $\delta = 12$ nm, $M_r\delta/H_c = 12.5$ nm, $k = 1$, and $g + t/2 = 45$ nm. It is clear that the most sensitive parameter is the magnetic spacing, followed by the read gap length, and then media parameters. If the areal density continues increasing at an annual compound rate of 60%, in fewer than five years, we will reach 1 Tbit/in^2 [268]. Now, the question is: What kind of sensor is required for this level of areal density?

To answer this question, we first look at what are the requirements on bit length and track width for terabit recording. Shown in Figure 13 are the estimated read track width and required sensor sensitivities at different areal densities by assuming an aspect ratio of 4:1. The solid and dashed curves are obtained by assuming the read track width to be 80 and 60% of the track pitch, respectively. The sensitivity requirement is obtained by assuming that the sensor will have to generate an output voltage of 1 mV. The read head for terabit recording would require a bit length as small as 12 nm, assuming that the read track width is 80% of the track pitch. This leads to a pulse width of about 36 nm, assuming that the channel density can be increased to 3. Assume that all other parameters are approaching their practical limits, such as $d = 5$ nm, $\delta = 5$ nm, and $M_r\delta/H_c = 6.25$ nm (maybe even smaller); then $g + t/2$ must be as small as 17 nm. This means that almost no read gap is allowed in sensors for terabit recording. This poses formidable challenges for CIP read sensors because one needs electrodes

between the two shields. Furthermore, the sensitivity of the sensor also has to be increased by a factor of 2–3, which means that the MR ratio must be around 30%. This might be possible for advanced spin valves at the sheet film level (as we have already achieved an MR ratio of about 27%), but it would be difficult to achieve the same figure at the sensor level.

Based on this background, the CPP spin valve is being studied as one of the possible candidates for high-density recording applications [269–279]. One of the advantages of the CPP spin valve over the existing Current-in-plane spin valve is that it allows the elimination of the insulating gap layer in the read head, which will contribute to the increase of the linear density [see Fig. 35(a)]. More importantly, the output voltage of the CPP sensor does not scale linearly with the width of the sensor as the CIP does. Assume that one has a read element with width W, height H, and thickness T, as shown schematically in Figure 35(b); the outputs of the sensor in the CIP and CPP modes are given by

$$\begin{aligned} \langle \Delta V_{\max}^{\mathrm{CIP}} \rangle &= \Delta\rho_{\mathrm{CIP}} \langle W \rangle J \\ \langle \Delta V_{\max}^{\mathrm{CPP}} \rangle &\approx \Delta\rho_{\mathrm{CPP}} \langle T \rangle J \end{aligned} \tag{31}$$

Here, $\Delta\rho_{\mathrm{CIP}}$ and $\Delta\rho_{\mathrm{CPP}}$ are the resistivity change due to the GMR effect (nominal values including the nonactive regions), $\langle W \rangle$ and $\langle T \rangle$ are the average width and thickness of the active region, and J is the current density. The above equations suggest that the CPP sensor is more suitable for ultrahigh-density recording because it is much easier to control the thickness than to control the lateral dimension. This is because the former is determined by the film thickness, while the latter is determined by lithography. Therefore, the CPP sensor is also advantageous from the manufacturability point of view. Note that the above discussion is only valid for constant current density operations. The dependence of the output signal on the sensor width will be relaxed for the CIP sensor when it is operated under a constant power.

5.4.2. Sensitivity Enhancement of CPP Sensors

Both theoretical and experimental studies on the CPP–GMR started shortly after the discovery of CIP–GMR in ferromagnetic/nonmagnetic multilayers [280–285]. It was found experimentally that the CPP–GMR is larger than the CIP–GMR in multilayer samples [280, 281, 286–288]. For spin valves, Vedyayev et al. estimated theoretically that the CPP–GMR should also be larger than the CIP–GMR [289]. By starting with the Boltzmann equation, Valet and Fert derived an analytical equation for the CPP–GMR [284]. This was later extended to the treatment of an exchange-biased spin valve with a typical structure AFM/FM/NM/FM [290, 291]. In the case where the spin diffusion length is much longer than the layer thicknesses, Bass and Pratt obtained [291]

$$A\Delta R = \frac{4(\beta\rho_F^* t_F + \gamma AR_{F/N}^*)^2}{2\rho_F^* t_F + \rho_N t_N + 2AR_{F/N}^* + AR_{AF/F}} \tag{32}$$

where t_F, t_N, and t_{AF} are the thicknesses of the ferromagnetic, nonmagnetic, and antiferromagnetic layers, and $AR_{AF/F}$ is the specific resistances at the anitiferromagnetic layer (AF)/ferromagnetic layer (F) interface. The bulk spin asymmetry coefficient (β), interface spin asymmetry coefficient (γ), ρ_F^*, and $AR_{F/N}^*$ are defined as

$$\begin{aligned} \beta &= \frac{\rho_F^{\downarrow} - \rho_F^{\uparrow}}{\rho_F^{\downarrow} + \rho_F^{\uparrow}}, & \gamma &= \frac{AR_{F/N}^{\downarrow} - AR_{F/N}^{\uparrow}}{AR_{F/N}^{\downarrow} + AR_{F/N}^{\uparrow}} \\ \rho_F^* &= \frac{(\rho_F^{\downarrow} + \rho_F^{\uparrow})}{4}, & AR_{F/N}^* &= \frac{(AR_{F/N}^{\downarrow} + AR_{F/N}^{\uparrow})}{4} \end{aligned} \tag{33}$$

where $\rho_F^{\uparrow}(\rho_F^{\downarrow})$ and $AR_{F/N}^{\uparrow}(AR_{F/N}^{\downarrow})$ are the bulk resistivity, and specific resistances at the F/N interfaces, for the majority and minority electrons, respectively. The above equation was obtained by assuming that the specific resistances between the ferromagnetic layer and antiferromagnetic layer or the contacts are independent of the direction of electron spin with respect to the local magnetization direction. On the other hand, in another extreme case where the spin diffusion length of the ferromagnetic layer is shorter than its thickness, $A\Delta R$ is given by [291]

$$A\Delta R = \frac{4(\beta\rho_F^* l_{sf}^F + \gamma AR_{F/N}^*)^2}{2\rho_F^* l_{sf}^F + 2AR_{F/N}^* + \rho_N t_N} \tag{34}$$

where l_{sf}^F is the spin diffusion length of the F layers, and all other notations are the same as those in Eq. (32).

When being used as a read sensor, the output voltage of a CPP SV sensor is given by $\Delta V = \alpha J A \Delta R$. Here, J is the current density whose maximum value is limited to $\sim 10^7$ A $\cdot$ cm^{-2} due to electron migration and power consumption considerations, and α is the head efficiency coefficient multiplied by the dynamic range coefficient of the sensor, which is about 0.2–0.5, depending on the structure of the read head. Assuming $\alpha = 0.2$ and $J = 5 \times 10^7$ A $\cdot$ cm^{-2}, an $A\Delta R$ of 10 m$\Omega \cdot \mu$m^2 is required to obtain an output voltage of 1 mV. A typical spin valve, however, can only have an $A\Delta R$ of about 1–2 m$\Omega \cdot \mu$m^2 [270, 274, 279]. Therefore, the performance of the CPP SV has to be improved before it can be applied to read heads. From Eq. (32), which applies to most of the practical cases, one can see that $A\Delta R$ can be increased by increasing the parameters in the numerator or decreasing the parameters in the denominator. We focus on the numerator first. The most straightforward way to increase $A\Delta R$ is to increase the thickness of the F layers. However, this is not a preferred approach because the free layer of a read sensor must

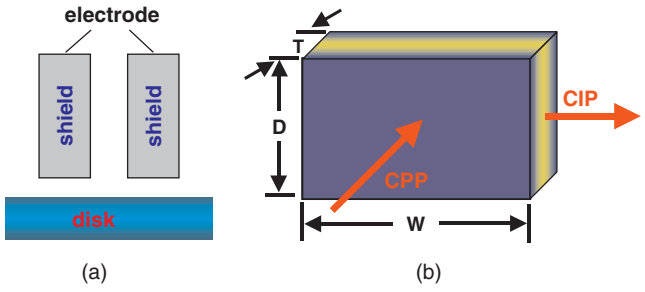


Figure 35. (a) Schematic drawing of CPP read head in which the shields can also be used as the electrodes which help to eliminate the read gap. (b) Schematic of a sensor element used in both the CPP and CIP modes.

be made as thin as possible; otherwise, it will degrade the sensitivity of the sensor. Although the values of other parameters, β, γ, and $AR^*_{F/N}$, can also be increased, their contribution to the enhancement of $A\Delta R$ is moderate because all of these parameters are determined by the properties of the constituent materials [272]. We are now left with only one parameter in the numerator, ρ^*_F, which is given by $\rho_F/(1-\beta^2)$. If one can find a way to increase the resistivity of the F layers without changing β significantly, then $A\Delta R$ will increase accordingly. It would be even more desirable should β be increased simultaneously. One of the possible approaches is to add nanooxides inside the F layers [270, 274]. The nanooxide effect is twofold. One of the effects is to increase the effective resistivity of the F layers without significantly changing the spin asymmetry coefficient. The second effect is to enhance the specular reflection of electrons so as to increase the number of F/N layers effectively. Nagasaka et al. showed that the nanooxide-added CPP spin valves exhibit an $A\Delta R$ 25 times larger than that of the conventional spin valves [274]. Focusing on the denominator, Gu et al. showed that the CPP–GMR in permalloy-based spin valves can be enhanced by increasing the spin-memory loss in ferromagnetic layers [292]. The particular structure that has been investigated by Gu et al. is Nb250/Cu10/FeMn8/Py12/Cu20/Py($t1$)/FeMn($t2$)/Py(12-$t1$)/Cu10/Nb250, where Py stands for permalloy, layer thicknesses are in nanometers, $t2 = 0.5$ or 1 nm, and $t1$ was adjustable from 1 to 11 nm. The ultrathin FeMn layer inside the free layer was used to increase the spin memory loss due to its extremely short spin diffusion length of about 1 nm. Although significant improvement in $A\Delta R$ has been obtained using this technique, the maximum value of $A\Delta R$ obtained is still below 2 $m\Omega \cdot \mu m^2$.

Among all of these techniques that have been investigated, the nanooxide gives the most significant enhancement to $A\Delta R$ without the need to introduce more layers or to increase the layer thickness physically; thus, it is desirable for practical applications. However, the oxide must be sufficiently thin; otherwise, it will affect the flow of electrons across the spin-valve stack in the CPP mode or increase the total resistance to an unacceptable level. However, if the oxide layer is too thin, it will lose its original functionality as an electron-reflection or resistivity-enhancement layer. Therefore, using oxide alone, one may not be able to obtain the same effect as what has been obtained in the CIP mode. In addition to the number of layers, $A\Delta R$ also increases with the bulk and interface spin asymmetry coefficient, as we discussed above. This might be achieved through using half metals as the electrodes, but so far, no suitable materials have been found. Based on this background, we have proposed a possible solution to overcome this difficulty [277].

It is known from the research work on semiconductor devices that electrons can be reflected not only by a single potential barrier of large height, but also by a superlattice-type of structure with moderate barrier heights at the interfaces, provided that the thickness of each layer is chosen such that it is one quarter of the de Broglie wave of electrons [293, 294]. This type of structure is often called an electron Bragg reflector [Fig. 36(a)]. It is relatively easy to realize this kind of structure in semiconductors because the de Broglie

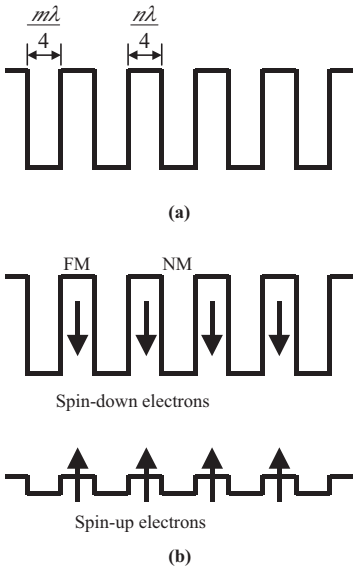


Figure 36. Schematic of a spin-independent Bragg reflector (a), and a spin-dependent Bragg reflector (b).

wavelength is relatively large. In metallic materials, however, the electron wavelength is close to the Fermi wavelength, which is just about 0.1–0.2 nm. If this is the "true" length scale, it is almost impossible to build a Bragg reflector for electrons using metallic materials. Fortunately, due to the discreteness of the atomic layers, the actual wavelength of the electron packet, or so-called envelope functions, is given by $2\pi/\Delta K$ [295]. Here, ΔK is the magnitude of the wavevector that spans the appropriate stationary points of the Fermi surface. This gives a wavelength of about 1.1 nm for Cu, for example. The Bragg reflection will occur when the conditions $d_{FM} = m\Lambda_{FM}/4$ and $d_{NM} = n\Lambda_{NM}/4$ are satisfied. Here, d_{FM} and d_{NM} are the thickness of the ferromagnetic and nonmagnetic layers, respectively, and m and n are odd integer numbers. If one chooses m and n in the range of 3–4, the thickness of each layer should be well within the controllable range of the state-of-the-art sputtering systems.

Now, a question arises here: How can the electron flow in the CPP mode if one has a perfect Bragg reflector in the spin-valve stack? The point is that the quantum-well states only form when all of the ferromagnetic layers are ferromagnetically coupled, and they only exist for the minority electrons [295]. In other words, the Bragg reflector in a magnetic/nonmagnetic multiple layer is a spin-dependent reflector [see Fig. 36(b)]. It reflects only the minority electrons, but lets the majority electrons flow almost freely. This will effectively increase the bulk and interface spin asymmetry coefficient, and thus $A\Delta R$ of spin valves. Note that the picture given here is based on the ballistic transport of electrons instead of diffusive transport; thus, Eqs. (32) and (34) might not be directly applicable to this case.

Figure 37 illustrates some of the possible structures of CPP sensors using the electron Bragg reflectors, which consist of a number of pairs of CoFe and Cu. The thickness of Cu can be chosen such that it facilitates a strong parallel coupling between the adjacent ferromagnetic layers, but at the same time, it satisfies the Bragg reflection condition.

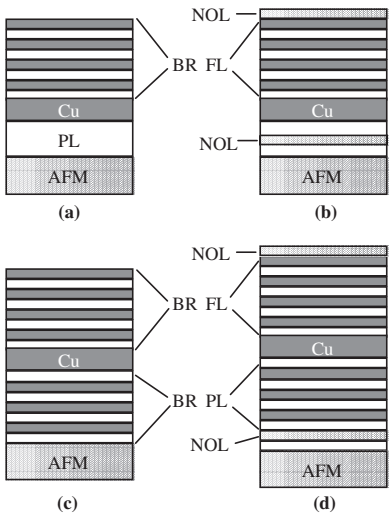


Figure 37. Schematic of some possible CPP spin valves using both nanooxide and Bragg reflectors. Here, BR stands for Bragg reflector, NOL refers to nanooxide, and FL (PL) is the free (pinned) layer.

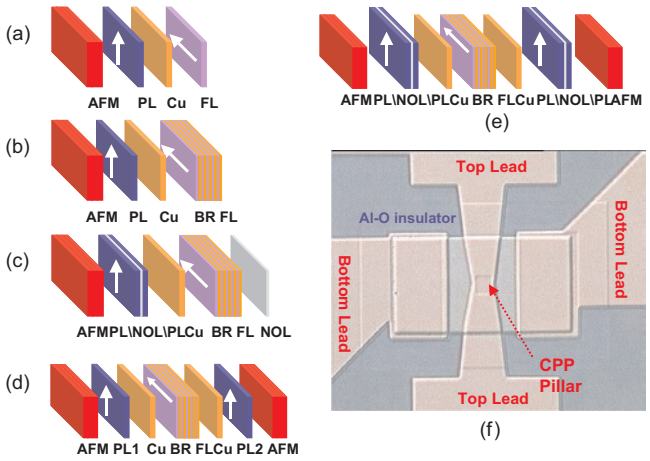


Figure 38. (a)–(e) Schematic drawings corresponding to sensor type A–E. (f) Optical micrograph of typical CPP sensors.

If the priority is to achieve ferromagnetic coupling, any difference in the thickness of the Cu layer to the value that is requested by the Bragg condition can be compensated by the thickness of the corresponding ferromagnetic layer. The m and n are not necessary to be the same; m can be small so as to reduce the total thickness of the Bragg reflectors. The Bragg reflectors can be used either in the free layer only [Fig. 37(a) and (b)] or in both the free and pinned layers [Fig. 37(c) and (d)]. To enhance the specular reflection effect, ultrathin oxides can be added at both the top of the free layer and the middle of the pinned layer [Fig. 37(b)]. The same approach can be taken for spin valves using Bragg reflectors at both the free and pinned layers [Fig. 37(d)]. Further to these steps naturally will be the one that adopts dual spin-valve structures, as is the case with Current-in-plane spin valves. The functions of the nanooxide are twofold. First, it increases the specular reflection. Second, it reduces the interface roughness so that the electron interference effect can occur.

A series of experiments have been carried out to verify the concept. The sheet films were fabricated using ultrahigh-vacuum sputtering. The detailed sputtering conditions can be found in [249, 250]. After the sheet film was fabricated, it was patterned into elements with different sizes using the e-beam lithography and ion-milling processes. The CPP spin valves that have been fabricated include: (1) simple single spin valves without nanooxides and Bragg reflectors (type A), (2) single spin valves with a Bragg reflector free layer (type B), (3) single spin valves with both a Bragg reflector free layer and a nanooxide cap layer (type C), (4) dual spin valves with a Bragg reflector free layer (type D), and (5) dual spin valves with both a Bragg reflector free layer and nanooxides at both the top and bottom pinned layers (type E). Figure 38 shows a schematic drawing of the spin-valve structures. The Bragg reflector used has a structure [CoFe(1.4 nm)/Cu(1.1 nm)] $\times$ 5. The antiferromagnet and pinned layers are IrMn (8 nm) and CoFe (3 nm), respectively. The top oxide was formed through the oxidation of NiFe, and the oxide in the middle of the pinned layer was formed through natural oxidation of CoFe. The total thickness of the CoFe pinned layer was always kept constant.

To ensure a reliable measurement of the change in the resistance–area product, we first tried to determine the dependence of the junction resistance on the junction area. Figure 39 shows the typical results for types A, B, C, and E spin valves. As can be seen from the figure, the junction resistance is inversely proportional to the junction area, indicating that current distribution is quite uniform over the entire junction. Similar results have been obtained for type D sensors. Table 3 lists the R_J and ΔR_J of the five different types of spin valves. In general, we can see that both the nanooxide and the Bragg reflector can improve the performance of the spin valve. The large junction resistance is due to the contact resistance at the sensor/electrode interface. We are currently developing a new technique to remove the junction resistance from the measured data. But

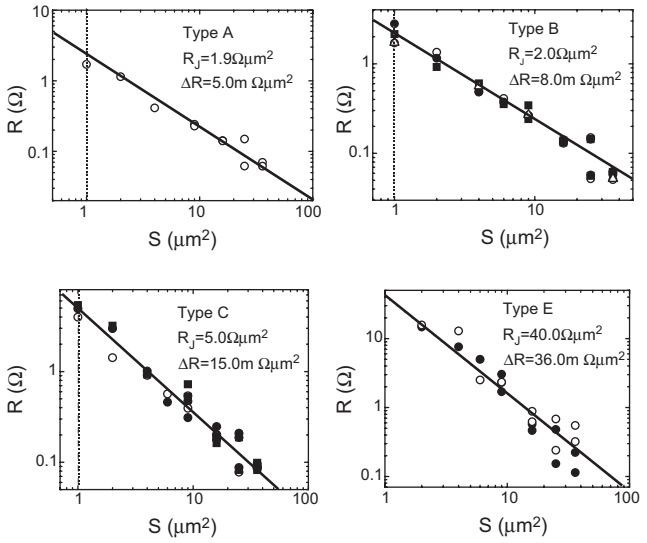


Figure 39. Dependence of the junction resistance on the junction area for different types of CPP spin valves.

Table 3. R_J and ΔR_J of different types of CPP spin valves.

	Type A	Type B	Type C	Type D	Type E
R_J ($\Omega \cdot \mu m^2$)	1.9	2.0	5.0	20	40
ΔR_J ($m\Omega \cdot \mu m^2$)	5.0	8.0	15	26	36

the junction resistance will not affect the accuracy of the measured $A\Delta R$ as much because we have already shown in Figure 39 that the current distribution problem is negligible in these samples.

The above results show that a well-designed and fabricated Bragg reflector may play the same role as a nano-oxide, but the former is more suitable for CPP spin valves because it does not increase the resistance substantially, as the oxides do. The results shown here are preliminary; we cannot exclude other effects introduced by the insertion of the laminated free layers. But we believe that this type of structure will become more useful in the future when suitable magnetic semiconductors are available because it is much easier to form Bragg reflectors using semiconductors. Although it is not in spin valves, a similar type of electron interference effect has already been observed in magnetic tunnel junctions grown epitaxially on a single-crystal substrate [296].

6. MAGNETIC TUNNEL JUNCTIONS

Another type of sensor that operates in the CPP mode is the magnetic tunnel junction (MTJ) [26–29]. The key difference between the CPP GMR sensor and the MTJ is that the electrical conduction in the former case is based on the spin-dependent scattering effect, both inside the ferromagnetic layers and at the ferromagnetic/nonmagnetic interfaces, while that in the latter is based on spin-dependent quantum mechanical tunneling across a thin potential barrier. To have a sound understanding of MTJ, we first look at the basic properties of a normal nonmagnetic tunnel junction. A typical quantum-mechanical tunnel junction is a sandwich of two conductive electrodes separated by a thin insulating barrier layer. For a tunnel junction with a sufficiently high potential barrier, the tunneling current can be calculated using the transfer Hamiltonian approach, which reads [297, 298]

$$I = A\int_{-\infty}^{\infty} |M|^2 D_1(E+E_{F1}) D_2 \times (E+E_{F2}+eV) \times [f(E+E_{F1}) - f(E+E_{F2}+eV)]dE \quad (35)$$

where

$$M = \frac{\hbar^2}{2m}\iint_S (\Psi_1 \nabla \Psi_2^* - \Psi_2^* \nabla \Psi_1)\cdot dS$$

is the matrix element of the transfer Hamiltonian, D_1 and D_2 are the density of states for the left and right electrodes, respectively, f is the Fermi distribution function, S is the cross-section area of the junction, m is the electron mass, $\hbar$ is the Planck constant, V is the bias voltage, E is the electron energy, E_{F1} (E_{F2}) is the Fermi level of the left (right) electrode, and A is a constant. Ψ_1 and Ψ_2 are the wavefunctions of the electrons at the initial and final states. In the case where the voltage applied across the junction is small, one has

$$f(E+E_{F_1}) - f(E+E_{F_2}+eV) = -\left.\frac{\partial f}{\partial E}\right|_{E=E_F} \cdot (eV) \quad (36)$$

$$G(V=0) = \frac{dI}{dV} \propto |M|^2 D_1(E_F) D_2(E_F) \quad (37)$$

where G is the conductance. For tunnel junctions with ferromagnetic electrodes, Eq. (37) is still valid, except that D_1 and D_2 are now dependent on the spin polarization of electrons. This is known to result in a different tunneling conductance between the parallel and antiparallel alignment of the magnetizations of the two electrodes, as shown schematically in Figure 40. The total conductance in the parallel configuration is thus approximately given by

$$G^{\uparrow\uparrow}(V=0) \propto D_1^{\uparrow}(E_F) D_2^{\uparrow}(E_F) + D_1^{\downarrow}(E_F) D_2^{\downarrow}(E_F) \quad (38)$$

and that in the antiparallel configuration is

$$G^{\uparrow\downarrow}(0) \propto D_1^{\uparrow}(E_F) D_2^{\downarrow}(E_F) + D_1^{\downarrow}(E_F) D_2^{\uparrow}(E_F) \quad (39)$$

Here, $D_1^{\uparrow}(D_1^{\downarrow})$ and $D_2^{\uparrow}(D_2^{\downarrow})$ are the density of states of the majority (minority) electrons of the two electrodes, respectively. In deriving the above equation, we have assumed that spin is conserved during tunneling. The junction magnetoresistance (JMR) is defined as

$$\text{JMR} = \frac{1/G^{\uparrow\downarrow}(0) - 1/G^{\uparrow\uparrow}(0)}{1/G^{\uparrow\downarrow}(0)} \quad (40)$$

which can be further reduced to

$$\text{JMR} = \frac{2P_1P_2}{1+P_1P_2} \quad (41)$$

Here,

$$P_1 = [D_1^{\uparrow}(E_F) - D_1^{\downarrow}(E_F)]/[D_1^{\uparrow}(E_F) + D_1^{\downarrow}(E_F)]$$
$$P_2 = [D_2^{\uparrow}(E_F) - D_2^{\downarrow}(E_F)]/[D_2^{\uparrow}(E_F) + D_2^{\downarrow}(E_F)] \quad (42)$$

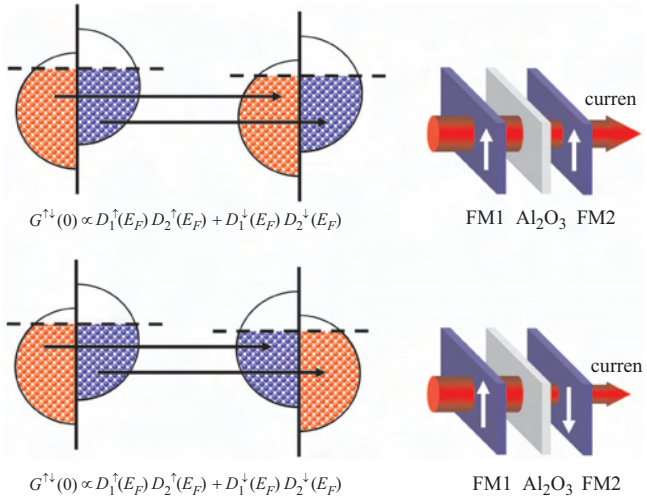


Figure 40. Schematic illustration of spin-dependent tunneling across an insulating barrier.

are the spin polarizations of electrons in the two ferromagnetic electrodes. This is the result that has been predicted by the Jullière model [26]. Assuming a typical value of 0.5 for P_1 and P_2, this simple model predicts a JMR of 40%. Note that an alternative definition for the tunnel magnetoresistance (TMR) frequently used in the literature is TMR $= 2P_1P_2/(1 - P_1P_2)$. Obviously, the TMR is always larger than the JMR, although the actual size of the magnetoresistance effect is the same. Although the Jullière model predicts a quite large MR, it was not until 1995 that MR ratios of about 20% were first reported by two groups for magnetic tunnel junctions at room temperature [28, 29]. Miyazaki and Tezuka [28] reported on an $Fe/Al_2O_3/Fe$ tunnel junction, 1×1 mm^2 in area, that exhibited a TMR ratio of 18% at room temperature. The barrier layer was formed through natural oxidation of 5.5 nm Al in air for 24 h. At about the same time, Moodera et al. fabricated magnetic tunnel junctions with CoFe and Co or NiFe as electrodes and oxygen-plasma oxidized Al as the barrier [29]. An MR ratio of 18% (in this case, JMR) has been obtained at low-bias voltage and room temperature in $CoFe/Al_2O_3/Co$ junctions with an area of 6×10^{-4} cm^2. The large MR ratio was attributed to the smoothness and small thickness of the electrodes. In both works, the magnetic tunnel effect was observed in large-area $FM/Al_2O_3/FM$ junctions fabricated by shadow masks. Gallagher et al. reported on the fabrication of microstructured magnetic tunnel junctions using a simple self-aligned lithographic process adopting an exchange-biased structure [299]. The combination of photo and electron-beam lithographies made it possible to cover a range of junction areas spanning five orders of magnitude (10^{-2}–10^3 μm^3). The MR ratio obtained ranges from 15 to 22% at room temperature and in fields of a few tens of Oersteds. The microstructured and exchange-biased magnetic tunnel junctions were further developed to have a higher MR ratio (42%) and moderate specific junction resistance ($\sim 60\ \Omega \cdot \mu$m^2), and were applied to prototypes of MRAMs [10]. After these pioneer works, a tremendous amount of effort was devoted to the development of magnetic tunnel junctions and their application in MRAMs and read heads. In what follows, we try to summarize some of the key issues that have been addressed in the literature for magnetic tunnel junctions.

6.1. Electrodes

6.1.1. Spin Polarization

The simple theoretical model suggests that ferromagnetic materials with large spin polarizations are the key to achieving a high MR. For a transition metal ferromagnet, the spin polarization is largely determined by whether the electrons participating in electrical conduction have an *s*-like characteristic or a *d*-like characteristic, or a mixture of these two. As the MR is determined by the spin-dependent density of states at the Fermi surface, its value is largely determined by the extent to which the *s* and *d* bands appear across the Fermi surface. However, the experimental determination of spin polarization for each material is quite challenging. The most straightforward way to determine the polarization is to use the spin-polarized photoemission spectroscopy technique, but it lacks the necessary energy resolution. An alternative technique that has been pioneered by Tedrow and Moservey is based on the spin-polarized tunneling across an FM-superconductor tunnel junction which gives a submillielectron volt energy resolution [300]. Recently, a new technique based on the point contact between the sample and a superconductor has been developed [301, 302]. This technique measures the polarization ratio using Andrew reflection. The point contact setup allows the measurement of samples in various forms, and it also does not require the use of a magnetic field. The values of P measured for some of the magnetic materials using both the point-contact Andrew reflection technique and tunneling techniques [301–305] are shown in Table 4 (note that spin polarization of Co and Ni should be negative in the processes where 3*d*-band electrons are dominant). Among them, NiMnSb, $La_{0.7}Sr_{0.3}MnO_3$, and CrO_2 are so-called half metals, which theoretically should have a polarization of 100%. As shown in Table 4, most soft magnetic materials have a polarization of 40–50%, corresponding to a JMR of about 30–40% (or a TMR of 38–67%). Theoretically, the sign of JMR can either be positive or negative, depending on the signs of P_1 and P_2. However, most of the experiments using NiFe, CoFe, Co, and Fe as the electrodes and Al_2O_3 as the barrier have produced positive MRs. Sharma et al. observed a negative MR in $NiFe/Al_2O_3/Ta_2O_5/NiFe$ composite barrier junctions, which was attributed to the different spin polarizations of NiFe at the two different interfaces, one with Al_2O_3 and the other with Ta_2O_5 [306]. Negative MR has also been observed by De Teresa et al. in $La_{0.7}Sr_{0.3}MnO_3$ (LSMO)/barrier/Co tunnel junctions, when $SrTiO_3$ (STO) was used as the insulating barrier [307]. This is in good agreement with the fact that the spin polarization of LSMO is positive [308], while that of the 3*d*-band density of states of Co at the Fermi level is negative. However, when Al_2O_3 was used as a barrier, the MR turned back to positive, and with a very different bias dependence. The authors concluded that the electronic structure of the barrier and the barrier–electrode interface has a strong influence on the spin polarization of electrons. Similar results have also been observed for magnetic tunnel junctions using $Ni_{40}Fe_{60}$ and Fe as the electrodes and STO as the insulating barrier [309], and in Fe_3O_4/STO/LSMO [310]

Table 4. Measured values of polarizations for different materials.

Materials	Polarization (%)	Ref.
Co	42 ± 2	[301]
	37 ± 2	[302]
Ni	46.5 ± 1	[301]
	32 ± 2	[302]
Fe	45 ± 2	[301]
$Ni_{80}Fe_{20}$	37 ± 5	[301]
$Co_{50}Fe_{50}$	$53\pm$	[301]
NiMnSb	58 ± 2.3	[301]
$La_{0.7}Sr_{0.3}MnO_3$	78 ± 4	[301]
$La_{2/3}Sr_{1/3}MnO_3$	95	[303]
CrO_2	90 ± 3.6	[301]
Fe_3O_4	-39	[304]
Sr_2FeMoO_6	60 ± 3	[305]

and CrO_2/natural barrier/Co [311] tunnel junctions. These results imply that the barrier material and/or the electronic structure of electrode/barrier interface play an important role in determining the sign of the spin polarization. This is consistent with the *ab initio* calculations of de Boer et al. on the Co(001)–HfO_2 interface [312], Oleinik et al. on the Co/Al_2O_3/Co interface [313], and Tsymbal et al. on the Fe/O surface [314], which suggest that the polarization of tunneling electrons is governed by the mechanism of electronic bonding at the metal/oxide interface.

6.1.2. Surface Roughness (Bottom Electrode)

One of the major problems that resulted in the failed attempts to produce high-MR MTJs before the early 1990s is related to the surface roughness of the first or bottom electrode on which the insulating barrier and top electrode are formed. If the surface roughness exceeds a certain critical value, the MTJ will fail either magnetically or electrically or in both ways. The former is mainly caused by the dipole or orange-peel coupling between the bottom and top FM electrodes, while the latter is caused by pinholes formed in the thin insulating barrier. Schrag et al. reported on measurements of the magnitude of Néel orange-peel coupling due to interface roughness in a series of magnetic tunneling junction devices [315]. The samples studied were Si(100) substrate/Ta/Al/NiFe/FeMn/Co(PL2)/Ru/Co(PL1)(pinned)/Al_2O_3(barrier)/NiFe(free)/Al/Ta with the thickness of the barrier, free, and pinned layers varying from sample to sample. Results from magnetometry and transport measurements are shown to be in good agreement with the theoretical model of Néel. In addition, the authors have also used transmission electron microscopy to directly probe the sample interface roughness, and obtain results consistent with the values obtained by magnetometry and transport methods. A similar type of study was also carried out by Li et al. using high-resolution transmission electron microscopy [316]. With an *in-situ* scanning tunneling microscope, Tegen et al. directly measured the roughness of the films, and found a close correspondence between the values for the coupling fields determined by the magnetooptical Kerr effect and the ones computed on the basis of the measured morphology parameters [317]. The authors confirmed an increase of the dipole coupling between the magnetic layers with decreasing barrier thickness, as predicted by the theoretical model [Eq. (25)]. In addition to topographic roughness, Tiusan et al. showed that, in samples involving polycrystalline magnetic films, beyond the orange-peel coupling, an important class of interaction is related to the dispersion fields associated with magnetic inhomogeneities, or magnetic roughness, arising from the local anisotropy fluctuations [318]. The roughness problem becomes more serious when fabricating MTJs for read-head applications due to the roughness of the lower shield layer. Sun et al. reported on the use of a gas cluster ion beam to smooth the shield layer, and successfully obtained an RA as low as 3.6–6.5 μm^2 and an MR ratio of 14–18% for MTJs grown on the smoothed shield layers [319].

6.1.3. Effect of Nonmagnetic Interfacial Layers

A spin-dependent quantum well forms in an ultrathin nonmagnetic metal sandwiched either from both sides by magnetic materials or from one side by a magnetic material and the other side by an insulator or vacuum [295, 320, 321]. In the former case, quantum-well states exist only when the magnetic layers are in parallel alignment, and in both cases, they are formed for minority electrons only. The observation of quantum-well states, however, requires atomically sharp interfaces due to the short wavelength of electrons in metals. For an ideal magnetic tunnel junction with a structure FM/NM/insulator/FM, theories predict an oscillation in the MR effect as a function of the NM layer thickness because the spin polarization of the tunneling electrons oscillates as a result of the resonant tunneling via the quantum-well states [322–324]. Moodera et al. fabricated Co/Au/Al_2O_3/$Ni_{80}Fe_{20}$ MTJs on liquid-nitrogen-cooled glass substrates with the thickness of Au varying from 0.1 to 1.2 nm [325]. It was found that the JMR decreases with the thickness of Au, which might be understood as the consequence of decrease of spin polarization [326–328]. But what was interesting was that a negative JMR was observed in the thickness range of 0.6–0.8 nm, which came along with an unusual bias dependence. The authors have shown that both results could be explained by taking into account the quantum-well states formed in the NM layer based on the Slonczewski model [329]. LeClair et al. also observed an inversion of sign of the MR in Co/Ru/Al_2O_3/Co tunnel junctions, but they attributed this to a strong density-of-states modification at the (interdiffused) Co/Ru interface [330].

Yuasa et al. fabricated MTJs on single-crystal Cu substrates of the structure Co(001)/Cu(001)/Al–O/$Ni_{80}Fe_{20}$ using molecular beam epitaxy, and they observed for the first time clear oscillation of the JMR as a function of the Cu layer thickness [296]. High-resolution transmission electron microscopy revealed that both the bottom electrode and the Cu layer were grown epitaxially on the Cu substrate, leading to atomically flat interfaces between Co and Cu and Cu and Al–O. As expected, the oscillation period at zero bias was about 1.1 nm, which is consistent with the value observed in GMR oscillation. As the spin polarization of electrons in Co and $Ni_{80}Fe_{20}$ was found to be positive in tunnel junctions with Al–O as the barrier, the negative peaks of the JMR observed in these junctions were attributed to the negative polarization of electrons formed in the Cu quantum wells.

6.2. Barrier Materials

6.2.1. Different Kinds of Barriers

The key to fabricating a high-quality MTJ is to form a pinhole-free and uniform ultrathin barrier layer of AlO_x [28, 29, 298, 331]. Although other types of insulating materials have also been investigated as the tunnel barrier, such as Ta_2O_5 [332–335], AlN_xO_y [336–339], $ZrAlO_x$ [340, 341], GaAs [342], GaO_x [343], diamond-like carbon [344], ZnS [345, 346], $SrTiO_3$ [347–350, 307, 351], MgF_2 [352], ZnSe [353], Cu_3N [354], MgO [355–359], HfO_x [360–362], TaO_x/AlO_x [306, 363], so far, AlO_x was found to be the most suitable barrier material for MTJ.

6.2.2. Formation and Characterization of AlO_x

The AlO_x layer can be formed using various methods, such as natural oxidation [28, 364–370], plasma oxidation [29, 371–377], ultraviolet (UV)-light-assisted oxidation [378–381], ion beam oxidation [382], radical oxidation [383], ozone oxidation [384], atomic layer chemical vapor deposition [385], and other equipment-specific techniques. In general, plasma or other energy beam-based techniques are faster and produce denser and more uniform oxides compared to natural oxidation, either in air or in pure oxygen. The property of the oxide formed is also dependent on the types of the process gases. Using $Kr–O_2$ plasma, Tsunoda et al. obtained a TMR of 58.8% at room temperature after annealing the junction Ta5/Cu20/Ta20/Ni–Fe5/Cu5/$Mn_{75}Ir_{25}$10/$Co_{70}Fe_{30}$2.5/Al–O/$Co_{70}Fe_{30}$2.5/Ni–Fe10/Cu20/Ta5 (thickness in nanometers) at 300 °C, while the achieved TMR ratio of the MTJ fabricated with the usual $Ar–O_2$ plasma remained 48.6% [386]. This improvement is remarkable considering the fact that the MR obtained so far using other oxidation methods is about 50% [387].

Under Oxidation of Al High-resolution transmission electron microscopy study by Bae et al. [388] revealed that natural oxidation occurred preferentially through the grain boundary of Al grains, leading to isotropically expanded AlO_x grains when Al is fully oxidized. On the other hand, in plasma oxidation, a flat AlO_x layer formed uniformly on the Al layer, leading to a sharp interface with the underlying metallic Al layer. Ando et al. conducted a systematic study of the microstructure of the barrier oxide formed using different methods [389]. It was found that it is easier to form a uniform oxide within a short time period using plasma oxidation. This is because, in addition to the full coverage of the Al surface, the energetic oxygen species are also able to penetrate into the grain boundaries. On the other hand, in the case of natural oxidation, it is difficult for oxygen to penetrate into the grain boundaries. Therefore, the as-deposited MTJs with barrier layers formed by plasma oxidation normally show a higher MR as compared to those with naturally oxidized barriers. Mitsuzuka et al. studied the interface structure of magnetic tunnel junctions with naturally oxidized AlO_x barriers using X-ray photoelectron spectroscopy (XPS) [390]. The MTJs studied had a structure Fe (50 nm)/AlO_x/CoFe (30 nm), where the barrier layer was formed through natural oxidation of Al with a thickness ranging from 0 to 5 nm. It was found that the MR showed a maximum value when Al is about 2–3 nm in thickness. The XPS analysis showed that an Al layer thicker than 1 nm covers the entire surface of the lower electrode. However, unoxidized Al remains when Al is thicker than 1 nm. For Al layers greater than 3 nm, the MR ratio is strongly affected by unoxidized Al, probably due to the decrease of spin polarization caused by the unoxidized Al layer. Zhang et al. reported on the effect of natural oxidation conditions on the performance of MTJs. It was found that natural oxidation at a lower oxygen pressure produces an MTJ with a lower RA and a smaller MR ratio as compared to those formed by a higher pressure oxygen oxidation [368]. Rutherford backscattering analysis confirmed the inadequate oxidation of Al at low oxygen pressure. Natural oxidation and UV-assisted oxidation are suitable for formation of MTJs of a low area–resistance product using ultrathin Al layers [391]. It was reported that a two-step or multiple steps of natural oxidation helps to improve the uniformity of the oxidized barrier layer [369, 392].

Pinholes The quality of the barrier layer plays a dominant role in determining the characteristics of the tunnel junction [393]. Allen et al. reported on the imaging of pinholes by electrochemical decoration of the pinholes using Cu [394]. The results obtained were found to be consistent with the breakdown voltage analysis [395]. Dimopoulos et al. reported that the "defective" sites of the barrier can also be detected using a barrier impedance scanning microscopy technique [396].

Inhomogeneity and Defects of the Barrier Layer Ando et al. studied the local current distribution of MTJ with a structure Ta5/$Ni_{80}Fe_{20}$5/IrMn15/Co5/Al1.3-oxidation using conducting atomic force microscopy [397]. Here, thickness is given in nanometers. It was found that the current distribution can be explained by assuming a Gaussian distribution for the barrier thickness with a mean value of 1.2 nm and a standard deviation of 0.1–0.15 nm. A similar study was also carried out by Luo et al. [398]. Rippard et al. used ballistic electron emission microscopy to study thin aluminum oxide tunnel junction barriers formed both by magnetron sputter deposition and thermal evaporation [399, 400]. It was found that the barriers made by oxidation of evaporated Al become fully formed at a thickness as small as 0.6–0.7 nm, while a thicker Al (0.9–1.1 nm) layer is necessary to form a continuous barrier by magnetron sputtering. Although the decrease of barrier thickness generally contributes to the reduction of RA, if the barrier is too thin, it will also cause reliability problems when the MTJs are used as read sensors or storage cell elements. Oliver et al. reported on dielectric breakdown studies on magnetic tunnel junctions having ultrathin barriers [401]. The MTJs, with a structure Ta5/PtMn25/CoFe2.2/Ru0.9/CoFe2.2/Al-oxidation/CoFe1/NiFe2.5/Ta15 (thicknesses are in nanometers), were fabricated by magnetron sputtering and patterned by deep ultraviolet photolithography to a junction area of about 0.2 μm^2. The tunnel magnetoresistance was 15–22% and RA 7–22 $\Omega \cdot \mu m^2$ for junctions having a 4.75–5.5 Å Al layer oxidized naturally. The authors observed two types of breakdown: abrupt dielectric breakdown at an effective field of 10 MV/cm determined by the thickness of the tunnel barrier, and a gradual breakdown related to defects in the tunnel barrier. After the breakdown, a metallic pinhole is created, the size of which depends on the maximum current applied to the junction. The current flowing through the pinhole creates a strong circular magnetic field that curls the local magnetization in the free layer around the pinhole. The subsequent free-layer reversal is very sensitive to the pinhole location. The electric properties after breakdown can be well described by an ohmic resistor and a tunnel magnetoresistor connected in parallel. Kikuchi et al. also reported similar types of phenomena in the breakdown processes [402]. The existence of pinholes in low-junction-resistance samples has been discussed in the context of a size-dependent breakdown voltage that decreases by increasing the junction size.

6.3. Annealing Effect and Thermal Stability

Postgrowth annealing at temperatures in the range of 220–400 °C was found to be effective in improving the performance of magnetic tunnel junctions, beyond which anneal will deteriorate the performances of MTJ [403]. The highest temperature at which the MR starts to decrease monotonically with temperature is highly dependent on the materials and device structures. This has a significant impact on the application of MTJs in MRAMs because standard back-end technology for the metallization of CMOS circuits requires annealing in the forming gas at 350–450 °C. Many attempts have been made to reveal the mechanisms for MR enhancement at intermediate temperatures and the loss of the MR at elevated temperatures. We summarize some of the results reported in the literature below.

6.3.1. Effect on the Barrier Layer

Anneal at moderate temperatures was found to improve the quality of the barrier layer through decreasing its thickness, increasing its barrier height, improving its homogeneity at interfaces with the electrodes, and reducing the defects inside the barrier. Parkin et al. reported on the thermal annealing study of an MTJ with a structure $Cr_{80}V_{20}25/Co_{75}Pt_{12}Cr_{13}15$/Al–plasma oxidation/Co15/Al20 or $Co_{88}Pt_{12}15$/Al20 [404]. Here, the layer thickness is in units of nanometers. It was found that the MTJ is thermally stable at temperatures greater than 300 °C. A comparison of cross-section transmission electron micrographs of an untreated sample and a similar one annealed at 350 °C indicates that the thickness of the amorphous tunnel barrier is slightly decreased after annealing. The resistance–area product and the MR of the devices initially increase slightly for lower temperature annealing treatments up to about 200–250 °C. For annealing at higher temperatures, both the MR and the resistance–area product decrease monotonically, although the MR decreases at a more rapid rate. MR values of ~10% were still obtained after annealing to ~350 °C. Due to the thick barriers used in these particular samples, the resistance–area product is in the range 5–6 $M\Omega \cdot \mu m^2$.

Sousa et al. reported on the thermal annealing effect of exchange-biased MTJs. Two series of samples were studied: one series with resistance–area products of 10–13 $M\Omega \cdot \mu m^2$, and the other with resistance–area products of 25–30 $k\Omega \cdot \mu m^2$ [405]. The former has a structure Ta7/Cu4/Ta7/NiFe6/CoFe3/Al_2O_3/CoFe3/MnRh18 or TbCo12/Ta3, and the latter has a structure Ta7/NiFe10/CoFe2/Al_2O_3/CoFe4/MnR17/Ta3. The barriers of samples with higher resistance–area products were formed depositing 1.8 nm Al, followed by 90 s plasma oxidation with a density of 6 mW/cm^2, and those of the lower resistance–area product samples were fabricated depositing 1.1–1.3 nm Al, and then followed, respectively, by 20 and 15 s plasma oxidation at a power density of 4 mW/cm^2. The junction tunneling magnetoresistance was found to increase from 22–26% in high-resistance samples and 22–37% in low-resistance samples, upon anneal up to 200–230 °C. Rutherford backscattering analysis suggested that anneal improved the asymmetry in both the oxygen distribution inside the barrier and the junction parameters. The MR increase in lower resistance samples was attributed to the increase of the barrier height. The increase of the barrier height and decrease of the barrier thickness were also observed by Nowak et al. [406].

Ando et al. studied the annealing effect on MTJs having the structure Ta/(Cu,Pt)/$Fe_{20}Ni_{80}$/IrMn/$Co_{75}Fe_{25}$/Al-oxide/$Co_{75}Fe_{25}$/$Fe_{20}Ni_{80}$/Ta [407]. When the Al thickness, oxidation time, and annealing temperature were 0.8 nm, 15 s (10 s), and 300 °C (250 °C), the TMR, RA, and TMR enhancement ratio against the as-grown samples were 49% (31%), 1.1 $k\Omega \cdot \mu m^2$ (230 $\Omega \cdot \mu m^2$), and 70% (100%), respectively. In order to investigate the annealing temperature dependence of the TMR ratio, the authors measured the local electrical properties for a Ta/$Fe_{20}Ni_{80}$/Pt/$Fe_{20}Ni_{80}$/IrMn/$Co_{75}Fe_{25}$/Al-oxide multilayer using the conductive atomic force microscopy technique. The current image became very homogeneous after annealing at around 300 °C for 1 h. The increase of the TMR ratio of the junction after annealing was thus attributed to both the increase of barrier height and the decrease of barrier height fluctuation. The degradation of TMR after annealing in excess of 350 °C was due to the decrease in barrier height, leading to the increase of leakage currents.

Schmalhorst et al. systematically investigated the structural, magnetic, magnetotransport, and tunneling properties of CoFe(1.5 nm)/Ru(0.9 nm)/CoFe(2.2 nm)/Al_2O_3/$Ni_{81}Fe_{19}$ (6 nm)/Ta (5 nm) junctions for different Al thicknesses and oxidation times after isochronal annealing up to 500 °C. The mean breakdown voltage of the junction increases with temperature, saturates at an annealing temperature of about 300–350 °C, and remains constant up to almost 500 °C. This implies that Al oxide is extremely stable, and its stability is improved by thermal annealing. The latter might be related to the defect reduction inside the barrier layer [408].

6.3.2. Effect on Electrodes

Dimopoulos et al. reported that, compared to Co/Ru/Co, Co/Ru/$Co_{50}Fe_{50}$ is more advantageous when being used as the pinned layer in MTJs due to its excellent thermal stability [409]. After successive annealing steps up to 400 °C, magnetic tunnel junctions with structures Cr (1.6 nm)/Fe (6 nm)/Cu (30 nm)/Co (1.8 nm)/Ru (0.8 nm)/$Co_{50}Fe_{50}$ (2.8 nm)/Al (1 nm)–plasma oxidation/CoFe (1 nm)/ Fe (6 nm)/Cu (5 nm)/Cr (3 nm) as pinned layers were found to still present a significant tunnel magnetoresistance of nearly 20%, and an almost intact magnetic rigidity of the hard magnetic system. The improvement in thermal stability was attributed to the improvement of the Ru/CoFe/AlO_x interfaces in terms of interdiffusion.

Due to the finite and negative enthalpies of formation of most transition metal oxides, the electrodes will also be partially oxidized during annealing, as revealed by X-ray photoelectron spectroscopy studies. However, the mechanism still remains unclear regarding how the partially oxidized interfacial layers are related to the enhancement or degradation of the tunnel magnetoresistance [410]. The oxidation of the electrode might be suppressed, if it has adverse influences on the MTJ, by inserting a thin layer of magnetic oxide at the barrier/electrode interface. During the annealing process, due to the much larger and negative enthalpy of

AlO_x as compared to most magnetic oxides and the larger affinity of oxygen for Al, oxygen will tend to move to the barrier layer instead of the electrode. Using this technique, Zhang et al. fabricated tunnel junctions with an interposed Fe oxide layer between the Al_2O_3 barrier (t_{Al} = 8–9 Å) and the top CoFe pinned layer, and obtained a large tunneling magnetoresistance of 39% after 40 min anneal at 380 °C [411, 412]. The annihilation of the CoFe oxide during postgrowth annealing was also observed by Dimopoulos et al. in Cr (1.8 nm)/Fe (6 nm)/Cu (30 nm)/Cr (1.8 nm)/Fe (6 nm)/Cu (30 nm)/$Co_{50}Fe_{50}$ (1 nm)/Al–plasma oxidation/Fe (6 nm)/Cu (10 nm)/Cr (5 nm) samples, leading to the improvement of aluminum oxide's stoichiometry [413]. Matsukawa et al. fabricated MTJs with structures of bottom electrode/PtMn/CoFe/Ru/ CoFe/$Fe_{1-x}Pt_x$/Al oxide/$Fe_{1-x}Pt_x$/NiFe/top electrode (x = 0.05–0.75), in areas from 2 × 2 to 30 × 30 μm^2, using conventional photolithography and the ion-milling method. After a postgrowth anneal at 400 and 420 °C, the TMRs of MTJs with $Fe_{1-x}Pt_x$ (x = 0.1–0.2) were still about 40 and 30%, respectively [414].

6.3.3. Diffusion of Mn

In addition to oxidation of electrodes, overoxidation of Al will also cause oxidation-induced diffusion of Mn from the pinning layer to the barrier layer during the postgrowth annealing. Samant et al. investigated the thermal stability of MTJs with a structure Si(100)/0.5 μm thick SiO_2/Ti5/Pd15/$Ir_{23}Mn_{77}$ or $Fe_{50}Mn_{50}$14/$Co_{80}Fe_{20}$2.4/Al1.3–plasma oxidation/$Co_{80}Fe_{20}$3/Pd1.5 (thickness is given in nanometers) using near-edge X-ray absorption fine-structure spectroscopy [415]. It was found that structures with IrMn antiferromagnetic exchange-bias layers are much more thermally stable than similar structures with FeMn exchange-bias layers. In either case, however, diffusion of Mn from the antiferromagnetic layer through thin exchange-biased ferromagnetic layers to the tunnel barrier is observed at elevated temperatures (>300 °C). Yoon et al. investigated the Mn diffusion in an NiFe/MnIr/CoFe/AlO_x multilayer after annealing at 300 °C using Auger electron spectroscopy and X-ray photoelectron spectroscopy, wherein it was found that the magnitude of Mn diffusion is correlated to the excess oxygen generated from the plasma oxidation of the AlO_x tunnel barrier [416]. The analysis showed that Mn diffusion was driven by the preferential oxidation of Mn at the oxide insulator interface due to its larger (negative) enthalpy of formation compared to iron and cobalt oxides. Saito et al. reported on the thermal stability study of dual-spin-valve-type double-tunnel junctions with structures Ni–Fe/Ir–Mn/$Co_{50}Fe_{50}$/AlO_x/$Co_{90}Fe_{10}$/AlO_x/$Co_{50}Fe_{50}$/Ir–Mn/Ni–Fe and Ni–Fe/Ir–Mn/$Co_{50}Fe_{50}$/AlO_x/$Co_{50}Fe_{50}$/Ni–Fe/$Co_{50}Fe_{50}$/AlO_x/$Co_{50}Fe_{50}$/Ir–Mn/Ni–Fe, fabricated using photolithography and ion-beam milling [417]. In order to clarify the mechanism of the loss of the MR ratio and that of $V_{1/2}$ (bias voltage at which the MR decreases to half of the zero-bias value) above 320 °C, the authors carried out an XPS study in $Ni_{81}Fe_{19}$/Ir–Mn/$Co_{50}Fe_{50}$/AlO_x multilayers annealed at various T_A. The Mn oxide peaks in addition to the Mn peaks (both $2p_{1/2}$ and $2p_{2/3}$ peaks were observed) were observed in the temperature range ≥300 °C. It was found that a small amount of Mn and Mn oxide reached the CoFe/AlO_x interface at T_A = 300 °C due to interdiffusion of Mn. The oxygen redistribution and homogenization occur between AlO_x and Mn in the range of T_A > 320–350 °C, leading to an increase of defect states in the barrier. This increase in defect states above 320–350 °C was assumed to be responsible for the decrease of both the MR ratio and $V_{1/2}$ due to the spin-independent two-step tunneling via defect states in the barrier. The Mn diffusion might be suppressed by increasing the pinned layer thickness [418].

6.4. Basic Characteristics of MTJ

A practical exchange-biased magnetic tunnel junction has a structure that is very similar to that of a spin valve, except that the Cu spacer is replaced by an ultrathin insulating layer (see inset of Fig. 41). Figure 41 shows a schematic of a typical MTJ structure. When measuring the magnetic response, a constant current is applied through one of the top and bottom electrodes, and the voltage drops across the other two terminals are measured when a magnetic field is swept within a certain range from a positive value to a negative value. However, caution must be taken when interpreting the data because the measured MR could be affected significantly by the current distribution effect [419–424].

Although the MR ratio is much higher than that of a typical spin valve, the MTJ does have a few drawbacks compared to spin valves when they are used as read sensors. These include: (1) large junction resistance, (2) bias voltage dependence of the MR, and (3) temperature dependence of MR. The junction RA of an MTJ can be as high as on the order of MΩ · μm^2. Although it can be lowered by reducing the barrier thickness and height, this will generally result in a lower MR ratio. For example, the MR ratio drops to about 13% when the RA is reduced to 5 Ω · μm^2 [425]. As an example, Figure 42 shows the magnetoresistance curve of a typical MTJ with a junction area of 80 μm × 80 μm fabricated in the author's group. For this specific device with a structure Ta20/$Ni_{81}Fe_{19}$20/$Co_{90}Fe_{10}$2/Al(1 nm)-oxidation/$Co_{90}Fe_{10}$2/$Ni_{81}Fe_{19}$7/$Fe_{50}Mn_{50}$20/Ta5 (the thicknesses of other layers are given in nanometers), the TMR ratio is 44.6% at a dc bias of 4.3 mV and the RA is about 250 kΩ · μm^2 [426]. Shown in Figure 43 is the bias voltage dependence of the MR ratio of this particular MTJ. As can be seen from the figure, the MR drops to almost half when the bias voltage is increased to about 0.4 V. Although this value does vary with the MTJ structure, it is roughly in the range of 0.3–0.5 V for most of the MTJ structures [427–435]. Although the mechanism is still not fully understood, it is believed that it is caused by both the bias-induced change in the barrier profile and the energy-dependent density of states for both the majority and minority electrons [436, 437]. It is argued that magnon excitation

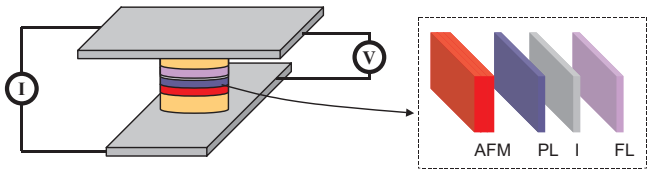


Figure 41. Typical structure of a magnetic tunnel junction device.

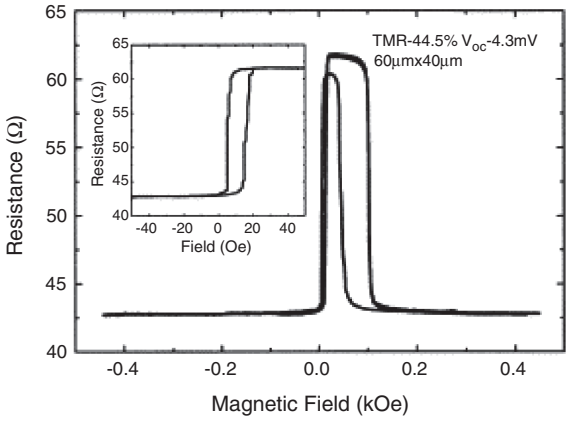


Figure 42. MR–H curve measured at room temperature for an MTJ with a structure Ta20/$Ni_{81}Fe_{19}$20/$Co_{90}Fe_{10}$2/Al(1 nm)-oxide/$Co_{90}Fe_{10}$2/$Ni_{81}Fe_{19}$7/$Fe_{50}Mn_{50}$20/Ta5. Inset shows the minor curve. Here, the thicknesses are given in nanometers. Reprinted with permission from [426], K. B. Li et al., *J. Magn. Magn. Mater.* 241, 89 (2002). © 2002, Elsevier.

in the magnetic layer might also be one of the possible reasons [438]. The energy-dependent density of states model can be roughly understood as follows. From Eq. (41), one can find that

$$\frac{\partial(\mathrm{JMR})}{\partial V} = \frac{2}{(1+P_1P_2)^2}\frac{\partial(P_1P_2)}{\partial V} \tag{43}$$

which reduces to $(\partial(\mathrm{JMR})/\partial V) = (4P/(1+P^2)^2)(\partial P/\partial V)$, when $P_1 = P_2 = P$. For electrodes with positive spin polarizations, the JMR will decrease with voltage when $(\partial P/\partial V) < 0$. Substituting Eq. (42) into $(\partial P/\partial V) < 0$, one obtains $(D^{\downarrow}(\partial D^{\uparrow}/\partial V) - D^{\uparrow}(\partial D^{\downarrow}/\partial V)) < 0$, which can be further simplified to $(\partial \ln(D^{\uparrow}/D^{\uparrow})/\partial V) < 0$. This implies that the JMR will drop if the ratio between the density of states for majority electrons and that of the minority electrons decreases with the bias voltage. As we will discuss shortly, the decrease of JMR with bias voltage is undesirable for practical applications. One of the possible approaches for reducing the bias voltage dependence is to use a double-tunnel junction [439]. Using the double junctions, it is possible to raise the voltage at which the MR drops to half of the low-bias value to 0.9 V, which is almost double that of a single-tunnel junction [440]. The voltage dependence can also be reduced using Hf-inserted AlO_x barrier layers [441].

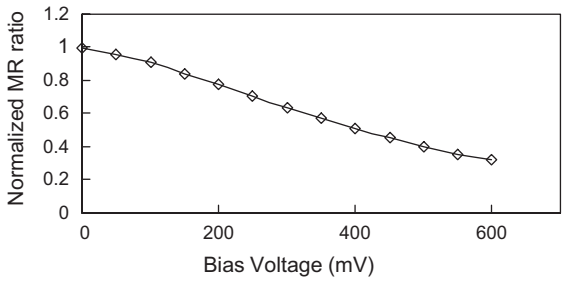


Figure 43. Bias voltage dependence of the MR ratio of an MTJ with the structure Ta20/$Ni_{81}Fe_{19}$20/$Co_{90}Fe_{10}$2/Al(1 nm)-oxide/$Co_{90}Fe_{10}$2/Ni_{81}Fe197/$Fe_{50}Mn_{50}$20/Ta5. Here, the thicknesses are given in nanometers.

In addition to voltage dependence, the MR ratio of MTJ was also found to decrease with temperature [430, 405, 442–446]. According to Shang et al. [442], Hagler et al. [444], and Wingbermühle et al. [445], the temperature-dependent total conductance of an MTJ can be written as

$$G = G_{DI}(T)[1 + P_1(T)P_2(T)\cos(\theta)] + G_{\mathrm{IN}}(T) \tag{44}$$

where $G_{DI}(T)$ indicates the direct tunneling conductance and $G_{\mathrm{IN}}(T)$ represents the spin-independent conductance via localized states, θ is the angle between the magnetizations of the two electrodes, and $P_1(T)$ and $P_2(T)$ are the temperature-dependent spin polarizations. The temperature dependences of $G_{DI}(T)$ and spin polarizations $[P_1(T)$ and $P_2(T)]$ are due to the temperature dependences of Fermi distribution and magnetizations, respectively, and are given by $G_{DI}(T) = CG_0T/\sin(CT)$ with $C = 1.39 \times 10^{-4}d/\sqrt{\phi}$, and $P_1(T) = P_{01}(1-\alpha_1T^{3/2})$, $P_2(T) = P_{02}(1-\alpha_2T^{3/2})$. Here, α_1 (α_2) is a material-dependent parameter, d is the barrier thickness in angstroms, ϕ is the barrier height in electronvolts, and G_0 is the value of $G_{DI}(T)$ at zero degrees Kelvin.

6.5. Applications of MTJ in Read Head

Magnetic tunnel junctions can be used either as read sensors in hard disk drives [447–453] or memory cells in MRAMs. Following are some advantages and drawbacks of MTJs compared to spin valves when they are used as read sensors.

6.5.1. Large Output Signal

For an MTJ sensor element with width W, height H, and thickness T, the maximum output voltage is given by

$$\Delta V_{\mathrm{MTJ}} = \frac{I\Delta RA}{W \cdot H} = J\Delta RA \tag{45}$$

where J is the current density, and ΔRA is the change in resistance–area product of the MTJ. The maximum output voltage refers to the value that is obtained when the magnetizations of the two FM layers are switched from a parallel alignment to an antiparallel alignment or vice versa. The actual output is normally a fraction of this due to the limited dynamic range of the read sensor and the head efficiency factor. In practice, however, the output is limited by the characteristic dependence of JMR on the dc bias voltage, that is [454],

$$\Delta V_{\mathrm{MTJ}}(V) = V_B \cdot \mathrm{MR}_0 \cdot (V_{\max} - V_B)/V_{\max} \tag{46}$$

where V_B is the bias voltage, MR_0 is the MR ratio at zero bias, and $V_{\max}$ is the voltage at which the MR becomes zero. In deriving the above equation, we have assumed that the MR ratio decreases linearly with the bias voltage. The maximum output signal is thus given by

$$\Delta V_{\mathrm{MTJ}}(V_B = V_{\max}/2) = \frac{1}{2}V_{\max} \cdot MR_0 \tag{47}$$

Assume that MR_0 = 5–10% at device level, and $V_{\max}$ = 800 mV; it gives an output voltage of 20–40 mV. On the other hand, the maximum output signal of a CIP spin valve

with thickness T, width W, current density J, and sheet resistance change ΔR_S is given by

$$\Delta V_{SV} = J \cdot \Delta R_S \cdot T \cdot W \tag{48}$$

Using typical values of $J = 1\times10^7$ A/cm^2, $\Delta R_S = 4\ \Omega$, $T =$ 30 nm, and $W = 1\ \mu$m, the maximum output voltage will be given by 12 mV. Of course, the actual output signal for both cases when being used as a read head is much smaller than these values due to the low head efficiency. We can see that the maximum output signal of an MTJ is much higher than that of a typical spin valve. Moreover, the output of the spin valve scales with its width. Therefore, the MTJ head becomes more advantageous than the CIP spin-valve head in terms of output signal when the track width shrinks into the submicron regime, which is already the case in the latest products.

6.5.2. Current-perpendicular-to-plane Operation

As is with the case of CPP GMR sensors, it is possible to remove the insulating gap in the MTJ heads, and thus they are more suitable for higher linear density heads.

6.5.3. Large Junction Resistance

The primary drawback of MTJ is its large junction resistance, in particular, when its size shrinks. The large junction resistance will result in both a small bandwidth and a low SNR of the read signal [455]. For high-frequency operations, the sensor resistance is required to be less than 100 Ω [425]. Figure 44 shows the dependence of the junction resistance as a function of the reader width by assuming that the sensor has a square shape and the resistance–area product is 1 $\Omega \cdot \mu$m^2. Also shown in the figure are the estimated sensor widths for applications at different recording densities. It is clear that the MTJ may not be suitable for read heads at an areal density of more than 200 Gbits/in^2. Jury and Wang reported on a possible approach to extend the bandwidth of the MTJ head using a simple buffer amplifier, which leads to an SNR improvement of 5–10 dB over a no-buffer configuration for frequencies between 100 MHz and 1 GHz, with a slight reduction of SNR at low frequencies [456].

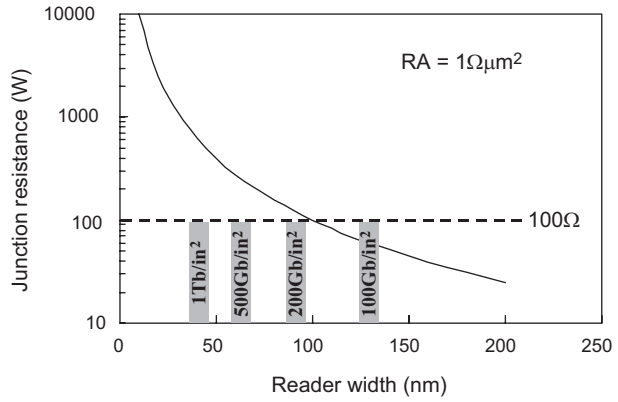


Figure 44. Junction resistance of MTJ as a function of reader width (see text for the assumptions that have been used to calculate the resistance).

6.5.4. Noises in MTJs

In addition to thermal noise, the MTJ also suffers from shot noise which increases with the sensing current [425, 455], $1/f$, and/or telegraph noise [457–461]. This makes the signal-to-noise ratio of MTJ smaller than that of the spin valve, even when the output signal is the same. The shot-noise voltage is given by $V_S = (2eVRB)^{1/2}$, while that of thermal noise is given by $V_{\text{th}} = (4k_BTRB)^{1/2}$. Here, e is the elementary charge of electrons, V is the bias voltage, R is the resistance of the sensor element, B is the bandwidth, k_B is the Boltzmann constant, and T is the temperature. Assuming $V = 400$ mV and $T = 300$ K, one can easily find that $V_S/V_{\text{th}} = 2.8$. The dominance of shot noise in the MTJ was found by Shimazawa et al. in their front flux guide-type magnetic tunnel junction heads [448]. George et al. reported that, depending on the quality of the barrier, the measured value of shot noise in an MTJ can be much lower than that predicted by theory [462]. In tunnel junctions with pinholes, the shot-noise voltage is reduced by a factor of $[R_P/(R_T + R_P)]^{1/2}$, with R_T and R_P the ideal tunnel resistance and the lumped parallel resistance due to the conductive paths of pinholes, respectively. Unfortunately, the decrease of shot noise due to pinholes does not necessarily lead to the increase of SNR because the pinholes also result in a decrease of JMR. Although it is not specific to MTJ, all types of MR heads also exhibit magnetization noise which may exceed the thermal noise when the sensor volume becomes very small [463].

Using the latest MTJ heads, Mao et al. [464] demonstrated an area density of 120 Gbits/in^2. Further optimization in materials, structures, and fabrication processes may improve the MTJ technology further, and render them suitable for higher areal density applications.

7. MAGNETIC RANDOM ACCESS MEMORY

7.1. Introduction

Although spin valves and MTJs have been invented for hard disk drive applications, they have also found important applications in memory devices. Computer memories featuring the following features are always highly demanded: (1) nonvolatility, (2) high density, (3) short cycle time, (4) low-power consumption, (5) low cost, (6) high reliability, (7) infinite lifetime, and (8) radiation hard (for special-purpose applications). A variety of memories have been commercialized. These include DRAM, SRAM, ROM, Flash EEPROM, and FeRAM. However, none of them is able to satisfy all of the requirements. Furthermore, most of these memories (e.g., DRAM) store information using charges; it is difficult to perform down-size scaling once it reaches a certain dimension due to the degradation of SNR. MRAM has been studied intensively as one of the promising candidates to replace some of the existing memories [9–11, 465–473]. In the case of disk drive applications, the spin valves and MTJs are employed as sensors to sense the magnetic fields from the information bits recorded on the hard disks. But when being used as cell elements in memories, the sensors themselves are used to store the information. The basic

structure of the memory cell is still the same as that of the sensor, but in the former case, one makes use of the saturation region of SVs and MTJs rather than the linear region used in read sensors.

7.2. Different Kinds of MRAM Designs

The concept of magnetic random memory itself is not new, but the early magnetic memory did not become popular due to both the drawbacks of its own design and the rapid advances made in semiconductor memories. However, the concept of using two crossed wires to write cells selectively is still applicable to current MRAMs. The key difference between the early magnetic memory and today's MRAM is that the former used an inductive signal to detect the storage state, while the latter utilizes AMR [9], GMR [474–479], or the spin-dependent tunneling phenomenon [10, 11, 299] to detect the signals. This has made it possible to remove the coils from the memory arrays, and at the same time, to increase the readout sensitivity, leading to the miniaturization of the cell elements. The first AMR-based MRAM was developed by Honeywell in the mid-1980s. However, its application was mainly limited to certain niche markets due to its high cost and relatively poor performance attributable to the low MR ratio. With recent advances in GMR and MTJ technologies, renewed interest has arisen in MRAMs since the mid-1990s. Due to the large read/write current and a relatively low SNR of large cell arrays of GMR-based MRAMs, most of the recent works on MRAMs are based on the MTJ design.

Figure 45 shows a schematic of an MRAM design using MTJs and the corresponding cell structure using one MTJ and one transistor per cell. The memory cells are formed at the cross junctions of the bit and word lines. Each memory cell consists of an MTJ for storing the information and a transistor for reading the cell selectively. The transistor can also be replaced by either a diode [480, 467] or can even be removed from the cell, as will be discussed shortly. A typical MTJ cell consists of a pinning layer to pin the reference layer, a storage layer to store the information, and a thin insulator to separate the two magnetic layers. The signal readout is based on the spin-dependent tunneling resistance of the MTJ by supplying a small sensing current to a specific cell. As the sensing current is only used to detect the relative orientation of the magnetization of the two magnetic layers instead of being used to switch the magnetization, it can be kept at a relatively low level. This is in contrast to the readout process of MRAMs using pseudospin valves, wherein larger currents are needed to switch the magnetization of the free layer from one direction to the opposite so as to detect the state of the stored bit (the so-called dynamic readout process).

7.3. Write Operation of MRAM

The write operation can be understood by looking at how a single-domain particle switches under two orthogonal fields, as shown schematically in Figure 46. Assuming that the magnetization switching occurs via coherent rotation, or based

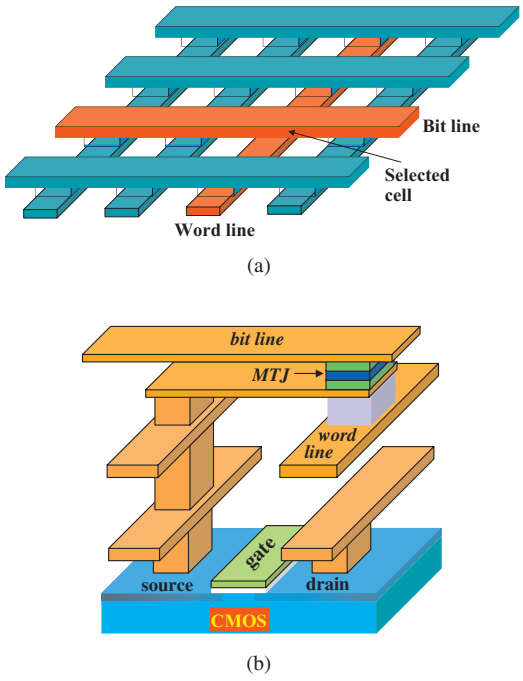


Figure 45. Schematic of the MRAM architecture using MTJ (a) and memory cell design using a transistor and an MTJ (b).

on the so-called Stoner–Wohlfarth (SW) model, the switching field can be obtained through the minimization of the energy density:

$$\varepsilon = -HM_s \cos(\phi - \theta) + K_u \sin^2 \theta$$

where H is the external magnetic field, M_s is saturation magnetization, K_u is the anisotropy energy constant, and θ and ϕ are the angle between the easy axis and M_s and H, respectively. Note that here only the Zeeman energy and the anisotropy energy are taken into account. Minimizing the energy density gives

$$h_\perp = -\cos^3 \theta, \quad h_{||} = \sin^3 \theta, \quad \text{or} \quad h_{||}^{2/3} - h_\perp^{2/3} = 1 \tag{49}$$

Here,

$$h_{||} = H_{||} M_s / 2K_u = \frac{H_{||}}{2K_u / M_s}$$
$$h_\perp = H_\perp M_s / 2K_u = \frac{H_\perp}{2K_u / M_s} \tag{50}$$

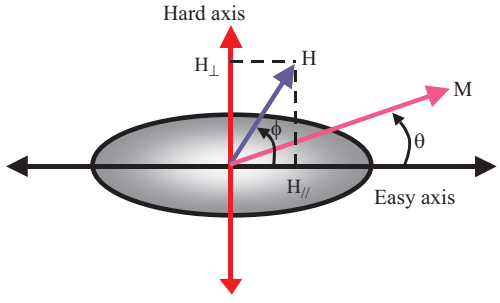


Figure 46. Schematic of the SW model for a single-domain particle.

are the reduced parallel and perpendicular components of the external field. Equation (44) gives an asteroid curve, as shown in Figure 47. To achieve magnetization reversal, the magnetic field vector must be such that its endpoint ends outside the asteroid (point *B*, for example). For some of the points inside the asteroid, although one of the field components is larger than that of point *B*, their magnetizations still cannot be switched. This simple mechanism makes it possible to write the individual cells selectively [481, 482].

In practice, however, the magnetization reversal process of the memory element is much more complicated than what is predicated by the simple single-domain model [469]. The magnetization switching process of rectangular elements is usually unrepeatable due to the formation of edge domains at the flat ends of the element. Elements with tapered ends have been introduced to suppress the formation of the edge domains [483]. Although the switching process becomes relatively repeatable, the tapered ends result in an increase in the switching field. An excellent review of the micromagnetics of MRAM was given by Zhu and Zheng [484]. Using Lorentz transmission electron microscopy, Kirk et al. [485, 486] and Yi et al. [487] found that the switching field of small magnetic elements can either increase or decrease, depending on whether the ends are gently curved or exhibit sharp corners. In element arrays, the magnetostatic interactions among the elements were also found to affect the switching process of individual elements [488, 489]. A unique design that makes the switching process repeatable, and yet is more suitable for ultrahigh-density memory is the vertical magnetoresistive random access memory (VMRAM) [471, 476, 484]. The VMRAM cell consists of a ring-shaped CPP spin valve with two FM layers separated by a nonmagnetic layer. The formation of a stable flux-closure configuration of the magnetization makes the switching process more robust and predictable.

The formation of edge domains is also sensitive to process-related defects in the magnetic element. This was found to result in fluctuations of the switching field among different cells [10]. As one of the examples, Figure 48 shows

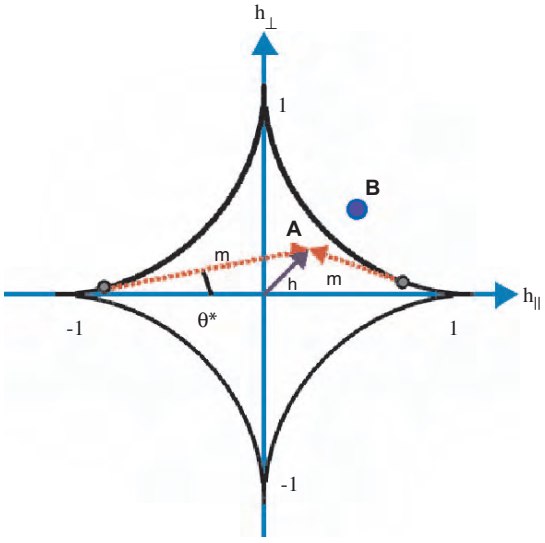


Figure 47. Asteroid curve for a single-domain particle.

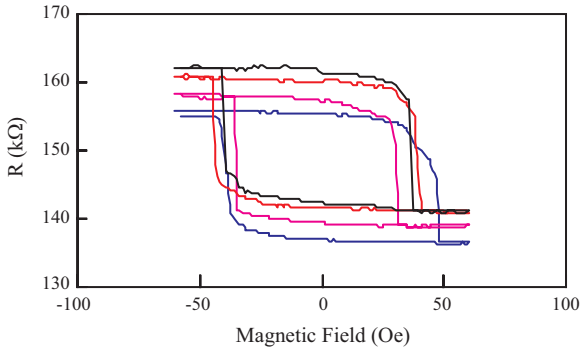


Figure 48. M–H curves of four identical cells with the structure Cr/Cu/Cr/Ta /NiFe /IrMn /CoFe/Ru/CoFe /AlO_x /CoFe /NiFe/CoFe/AlO_x/CoFe/Ru/CoFe/IrMn/Ta.

the *M–H* curve for a double-tunnel-junction MRAM with four cells of the same size fabricated in the author's group. Although they were fabricated using identical processes and were placed one next to another on the same wafer, the switching fields differ from one another. In addition to the fluctuation of switching fields among different cells, the shift of the switching curve along the field axis also requires a large tolerance for the switching current [490]. The latter stems from various magnetic couplings between the free and pinned layer, and thus must be minimized by process optimization. In the case of mass production, how to achieve a narrow distribution for the switching field for millions or billions of cells is a great challenge.

In addition to switching field fluctuation reduction, efforts have also been made to reduce the absolute value of the switching field, which is proportional to the write current. Currently, the write current is about a few milliamperes; it will increase to more than 10 mA when the element size shrinks to less than 100 nm [491]. Tehrani et al. reported that the effective write current can be reduced by covering the word line (Cu wire) with a high-susceptibility soft magnetic layer from all sides but the side that is facing the MTJ cell [492]. Under optimum conditions, it is possible to reduce the write current by a factor of 2–4 [491]. Further reduction of the write current can be achieved by reducing the distance between the word line and the MTJ cells. When the demagnetizing field is taken into account, the switching field of an MRAM cell using MTJ is proportional to the thickness of the free layer and inversely proportional to the size of the cell. Therefore, it is important to keep the thickness/width ratio constant during down-size scaling of the cell. Finally, the write current can also be reduced by reducing the shape anisotropy and increasing the anisotropy of the material itself by using, for example, synthetic ferrimagnets as the free layer [493].

7.4. Read Operation of MRAM

As we mentioned in the beginning of this section, one of the key features of the MRAM is that the information is stored in the relative orientation of the magnetization of two separate magnetic layers rather than in the charged/discharged state of capacitors. Therefore, it is more suitable for scaling down to small structures, at least, in terms of output signal.

As discussed in the MTJ section, the maximum output signal of an MTJ is given by

$$\Delta V_{\mathrm{MTJ}}(V = V_{\max}/2) = \frac{1}{2} V_{\max} \cdot \mathrm{MR}_0 \qquad (51)$$

In principle, it is independent of the cell size, although it may decrease by reducing the size in actual devices. As the MTJs in MRAMs detect the resistance difference between parallel and antiparallel states, the peak-to-peak output voltage is in the range of 120–160 mV, which is much larger than the output voltage of a read sensor. The use of double-tunnel junctions may further boost this value [494]. One of the key concerns in the readout process of MRAMs at the product level is the fluctuation of MR values among the large number of cells. Fortunately, the rapid progress in process development has now made it possible to suppress the fluctuation of MR at an acceptable level. For example, Tehrani et al. reported that the mean value and standard deviation of MR in their particular samples were successfully suppressed to 5 and 1%, respectively, over a 200 mm wafer [492].

7.5. Diode-Free/Multilayer/Multilevel Designs

The existing 1 MTJ/1 T cell design not only makes the MRAM costly, but also makes it difficult to reduce the dimensions of the entire MRAM cell. A switch-free design is much more desirable for fully making use of the advantages of MRAMs [495, 496]. Figure 42 shows an example of such a MRAM design using MTJs. As there is no switching diode or transistor associated with each MTJ, all of the cells are electrically mutually connected, forming a large resistor network. This will reduce the amplitude of the readout signal due to the current-shunting effect, as shown schematically in Figure 49. The author's group has proposed a simple peripheral circuitry to reduce the shunting current [496]. As shown schematically in Figure 50, the simple circuit is basically a voltage follower. It sets an equal electrical potential for adjacent bit lines so as to prevent current shunting to other paths when a selected bit is being read. Although it is only shown here for a 2×2 cell, it can be easily scaled up to large-scale cell arrays without the need to increase the complexity of the peripheral circuit. The effectiveness of this

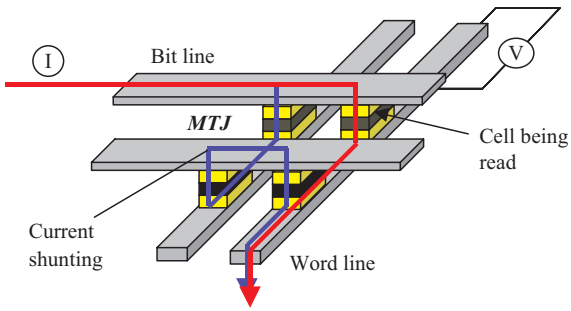


Figure 49. Schematic diagram of a switch-free 2×2 MRAM cell array. When a sensing current is supplied to a selected cell to read out the data, a part of the current is shunted to other cells, resulting in a decrease in the readout signal.

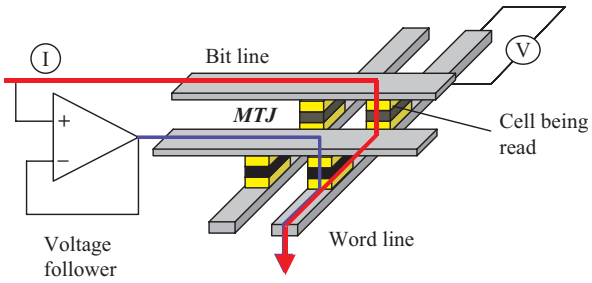


Figure 50. Schematic diagram of a switch-free 2×2 MRAM cell array. A simple peripheral circuitry is used to reduce the shunting current.

method has been verified experimentally [496]. The diode-free design is particularly useful for building multiple-layer MRAMs. In addition to multiple-layer design, multilevel or multistate design also remains as one of the options for realizing gigabit MRAMs in the future [497–500].

7.6. Electrical Field Driven Memory

The MRAM in its present form is intrinsically a current-driven device, at least during writing. The most desirable memory should be the one such that only reading involves current, while writing is based on an electrical field effect. This kind of memory is difficult to be implemented using metallic magnetic materials because their magnetic properties cannot be altered by a static electric field, at least at the moment and under normal conditions. The necessity of generating a magnetic field from a current to switch the magnetization in current MRAMs increases the complexity of the design and manufacture of such devices. Although magnetization switching based on "spin transfer" has been studied extensively, which has the potential to simplify the MRAM design, the current that is needed to switch the magnetization is still comparable to the write current of existing MRAMs [501–508]. Very recently, Zhuang et al. [509] demonstrated the first prototype of an electrical-field-driven RAM using the colossal magnetoresistance effect (CMR). As it is based on the huge change of resistance at two different states, it is also dubbed a resistance random-access memory or RRAM. For a $0.8\ \mu\mathrm{m} \times 0.8\ \mu\mathrm{m}$ cell, the write current was 0.2 mA, which is much smaller than the write current of MRAMs. The RRAM is particularly suitable for multilevel memories due to its large resistance change between two different states (by a factor of 10–1000). Another field which can be explored for creating electrical-field-driven memory devices is magnetic semiconductors. Ohno et al. recently demonstrated electrical-controlled ferromagnetism in (In, Mn)As-based magnetic semiconductors, but so far, it has not been implemented in any form of functional devices [510].

8. SEMICONDUCTOR SPINTRONICS

In this chapter, we have focused on metal-based spintronics and its application in data storage. Another important emerging field is the spintronics that uses nonmagnetic semiconductors, magnetic semiconductors, and metal/semiconductor hybrid structures. A number of device

concepts have been proposed, and some of them have been verified experimentally. These include, but are not limited to: (1) the spin field-effect transistor [511], (2) the spin-valve transistor [512], (3) the magnetic semiconductor tunnel junction [513, 514], (4) magnetic semiconductor diodes [515], (5) the hybrid Hall effect device [516], magnetic tunnel transistors [517], and spin-dependent light-emitting devices [518–520]. The greatest advantage of spintronics using semiconductors is that it can be integrated into the existing microelectronics. Furthermore, it also offers the possibility of controlling the magnetic properties using an electric field, which may lead to novel nanodevices, which is not possible using metal-based magnetic materials. Research activities in semiconductor spintronics have expanded almost exponentially in recent years. An in-depth coverage of this topic is out of the scope of this chapter because we focus on data storage, which at the moment is still dominated by metal-based spintronics. Semiconductor-based spintronic devices may find applications in data storage, but they need to overcome many technical hurdles before they can be used in hard disk drives. This is because most of the semiconductors can only be grown on high-quality crystalline substrates instead of the AlTiC substrates used currently for sliders. Most of the semiconductor epilayers are also too fragile to sustain the various lapping/polishing processes employed in current magnetic head fabrication processes. Therefore, semiconductor-based spintronic devices are more suitable for future memory and logic devices. The latest reviews of semiconductor-based spintronics can be found in [4, 31], and the April 2002 issue of *Semiconductor Science and Technology*.

9. FUTURE TREND OF DATA STORAGE AND THE ROLE OF SPINTRONICS

Most of the existing data storage devices can be categorized into three major groups: electronic, magnetic, and optical devices. The electronic storage devices store the information using charges. Therefore, the absolute number of electrons that are used to store one bit of information matters because it will determine the SNR of the readout signal. To this end, various types of capacitance enhancement techniques have been developed, such as bottle-shaped deep trenches, hemispherical grains, and three-dimensional structures over the bit line. In addition to the innovation in designs, efforts are also being made to develop suitable high-K capacitor dielectrics. In the field of flash EEPROMs, multibit and multilevel technologies are also being introduced to increase the storage capacity.

Optical storage, on the other hand, has the most diversity among all kinds of data storage devices. This is because, as shown schematically in Figure 51, light as an electromagnetic wave carries more information than its static counterpart of electrical and magnetic fields. Depending on which parameter in the electrical field expression that is used to read or store the information, various types of optical data storage technologies have been developed, and some of them have been commercialized. These include: (1) read-only memory (ROM), (2) phase-change (PC) disk, (3) magnetooptical (MO) disk, (4) multiple-layer storage, (5) holographic

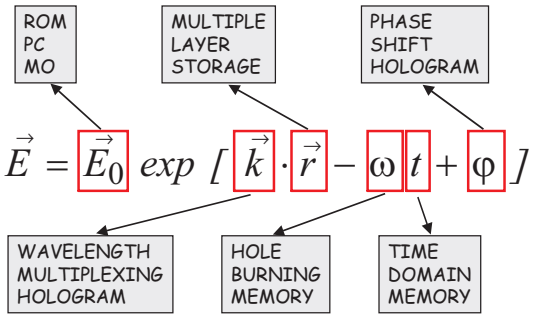


Figure 51. A wide range of optical data storage technologies that has been developed using the different properties of light waves.

data storage, (6) hole-burning memory, and (7) time-domain memory. The wave nature of light makes it possible to realize the storage of information in real space [(1)–(4)], reciprocal space [(5), (6)], and the time domain. If we focus only on real space recording, the areal density of optical storage is limited physically by the fact that a certain number of photons are required to reproduce the information with an acceptable SNR. When the bit becomes very small, the number of photons that come back from a tiny bit is limited by the intensity of the light, which is in turn determined by the damage threshold of the storage medium. Of course, in practice, the areal density of real space optical recording is largely determined by how small one can focus the laser spot into and, at far field, it is determined by the diffraction-limited performance of the optical system. Although near-field optics can reduce the spot size to a certain extent, it is still insufficient to compete with the state-of-the-art magnetic recording devices.

The magnetic storage devices which have been the focus of this chapter store the information using the spin of electrons. Although the output of CIP sensors scales with the size of the information bit, the output of CPP sensors including MTJ is relatively insensitive to the size of the sensor. However, even for the latter, the SNR scales inversely with the bit size. Therefore, the areal density of magnetic recording will be limited either by the SNR or by the thermal stability of the bit, whichever gives the lower density. Although large ballistic magnetoresistance has been obtained in magnetic nanocontacts [521, 522], one still has to find a way to convert it into a useful sensor for read heads, in particular, the linearity and low-field requirements.

If the bit size of all of these different types of storage devices keeps shrinking, it will eventually approach the size of atoms. When we reach that stage, we will probably have to seriously look at the information storage in reciprocal space. As mentioned above, the concept of reciprocal space recording has already been explored in optical data storage. A similar concept should also be explored in charge- and spin-based memories. This will lead to the general concept of quantum information storage. Nanometer-scale spintronics is expected to play a bridging role to bring the classic data storage into the quantum information storage era (see Fig. 52).

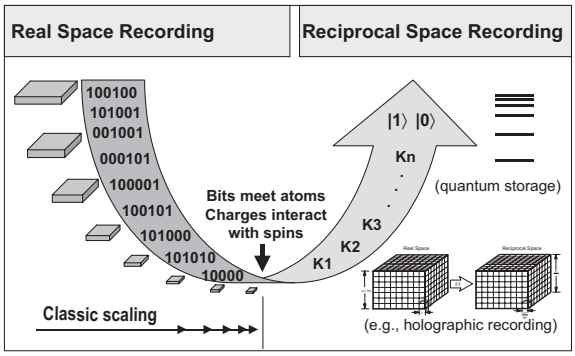


Figure 52. Future trend of data storage.

GLOSSARY

Anisotropic magnetoresistance (AMR) A phenomenon that describes the variation of the resistance of a single piece of magnetic material when it is subjected to an external field.

Current-in-plane GMR Magnetoresistance for the case that when the current is parallel to the film plan.

Current-perpendicular-to-plane GMR Magnetoresistance for the case that when the current is perpendicular to the film plan.

Giant magnetoresistance (GMR) A phenomenon that describes the variation of the resistance of a laminar structure of magnetic and nonmagnetic materials when it is subjected to an external field.

Magnetic data storage A data storage device or system that stores information by the relative orientation of the magnetization of localized regions of a storage medium.

Magnetic nanostructures Magnetic materials/structures with a size comparable to those of domain wall or Fermi wavelength of electrons in one or more spatial directions.

Magnetic random access memory (MRAM) A nonvolatile memory whose storage cell is a magnetic tunnel junction.

Magnetic semiconductor A semiconductor with ferromagnetism.

Magnetic tunnel junction (MTJ) A triple layer of magnet/insulator/magnet whose resistance is dependent on the relative orientation of the magnetizations of the two magnetic electrodes when electrons tunnel through the barrier.

Nanomagnetism Magnetism at nanometer scale.

Nanotechnology A frontier in the multidisciplinary field of physics, chemistry, biology, and engineering which allows one to design and fabricate materials, devices, and systems at atomic and molecular level.

Non-volatile memory A memory that maintains the information stored when the power is off within a certain time period.

Read sensor A device that is used to read the information from the storage medium by detecting the magnetic fluxes changes from the magnetic transitions.

Real space recording Information storage devices or systems which store information in real space.

Reciprocal space recording Information storage devices or systems which store information in K or energy space.

Spintronics Electronics that makes use of both the charge and spin degrees of freedom of electrons.

Spin-valve A triple layer consists of a hard magnetic layer, a non-magnetic spacer, and a soft magnetic layer whose resistance is dependent on the angle between the magnetization directions of the two magnetic layers. When being used at low fields, the magnetization of the hard layer is considered "fixed" while that of the soft layer is free to respond to the change of the external field.

ACKNOWLEDGMENTS

The author is grateful to the members of the Nano Spinelectronics Group at the Data Storage Institute, especially Dr. Kebin Li, Dr. Jinjun Qiu, Dr. Yuankai Zheng, Dr. Zaibing Guo, Dr. Guchang Han, and Ms. Luo Ping for their assistance with some of the experimental data. He is also grateful to Dr. Li Wang and Mr. Yatao Shen for their assistance with some of the literature survey work. He would also like to acknowledge the support of Prof. Chong Tow Chong, Dr. Thomas Liew, and Dr. Chang Quan Teck.

REFERENCES

1. G. A. Prinz, *Science* 282, 1660 (1998).
2. Special issue on magnetoelectronics, *Phys. Today* (Apr. 1995).
3. J. F. Gregg, I. Petej, E. Jouguelet, and C. Dennis, *J. Phys. D: Appl. Phys.* 35, R121 (2002).
4. M. Ziese and M. J. Thornton, Eds., "Spin Electronics." Springer, Berlin, 2001.
5. J. Hong, J. Kane, J. Hashimoto, M. Yamagishi, K. Noma, and H. Kanai, "12th Magnetic Recording Conference." Minneapolis, MN, Aug. 2001.
6. M. Yamagishi, J. Hong, E. N. Abarra, B. R. Acharya, I. Okamoto, and H. Komoriya, "Digest of INTERMAG Europe 2002," Amsterdam, 2002, paper BA01.
7. Z. Y. Zhang, Y. C. Feng, T. Clinton, G. Badran, N.-H. Yeh, E. Girt, S. Harkness, M. Muntteanu, H. J. Richter, R. Ranjan, S. Hwang, G. C. Rauch, M. Ghaly, D. Larson, E. Singleton, V. Vas'ko, J. Ho, F. Stageberg, V. Kong, K. Duxstad, and S. Slade, "Digest of INTERMAG Europe 2002," Amsterdam, 2002, paper BA02.
8. K. Stove, F. Liu, Y. Chen, X. Dang, P. Luo, J. Chen, K. Kung, M. Lederman, M. Re, G. Choe, and M. Zheng, "Abstract of the 47th Annual Conference on Magnetism & Magnetic Materials," Tampa, FL, Nov. 2002, paper AG08.
9. J. M. Dauton, *Thin Solid Films* 216, 162 (1992); J. M. Dauton, *J. Appl. Phys.* 81, 3758 (1997).
10. S. S. P. Parkin, K. P. Roche, M. G. Samant, P. M. Rice, R. B. Beyers, R. E. Scheurlein, E. J. O'Sullivan, S. L. Brown, J. Bucchigana, D. W. Abraham, Y. Lu, M. Rooks, P. L. Trouilloud, R. A. Wanner, and W. J. Gallagher, *J. Appl. Phys.* 85, 5828 (1999).
11. S. Tehrani, E. Chen, M. Durlam, M. Deherrera, J. M. Slaughter, J. Shi, and G. Kerszkowski, *J. Appl. Phys.* 85, 5822 (1999).
12. B. K. Middleton, in "Magnetic Recording Technology" (C. D. Mee and E. D. Daniel, Eds.), Chap. 2. McGraw-Hill, New York, 1996.
13. S. X. Wang and A. M. Taratorin, "Magnetic Information Storage Technology." Academic, New York, 1999.
14. R. M. White, "Introduction to Magnetic Recording." IEEE Press, New York, 1985.
15. H. N. Bertram, "Theory of Magnetic Recording." Cambridge University Press, Cambridge, 1994.
16. H. J. Richter, *J. Phys. D: Appl. Phys.* 32, R147 (1999).

17. E. N. Abarra, A. Inamota, H. Sato, I. Okamoto, and Y. Mizoshita, *Appl. Phys. Lett.* 77, 2581 (2000).
18. E. E. Fullerton, D. T. Margulies, M. E. Schabes, M. Carey, B. Gurney, A. Moser, M. Best, G. Zeltzer, K. Rubin, H. Rosen, and M. Doerner, *Appl. Phys. Lett.* 77, 3806 (2000).
19. E. N. Abarra, B. R. Acharya, A. Inomata, and I. Okamoto, *IEEE Trans. Magn.* 37, 1426 (2001).
20. R. Hunt, *IEEE Trans. Magn.* MAG-7, 150 (1971).
21. B. Gurney, M. Carey, C. Tsang, M. Williams, S. S. P. Parkin, R. Fontana, Jr., and E. Grochowski, *Datatech* 4, 87 (2000).
22. M. N. Baibich, J. M. Broto, A. Fert, F. Nguyen van Dau, F. Petroff, P. Etienne, G. Creuzet, A. Friederich, and J. Chazelas, *Phys. Rev. Lett.* 61, 2472 (1988).
23. G. Binasch, P. Grunberg, F. Saurenbach, and W. Zinn, *Phys. Rev. B* 39, 4828 (1989).
24. B. Dieny, V. S. Speriosu, S. S. P. Parkin, B. A. Gurney, D. R. Wilhoit, and D. Mauri, *Phys. Rev. B* 43, 1297 (1991).
25. A. Vedyayeva, M. Chshieva, N. Ryzhanovaa, B. Dieny, C. Cowacheb, and F. Brouersc, *J. Magn. Magn. Mater.* 172, 53 (1997).
26. M. Jullière, *Phys. Lett.* 54A, 225 (1975).
27. S. Maekawa and U. Gafvert, *IEEE Trans. Magn.* MAG-18, 707 (1982).
28. T. Miyazaki and N. Tezuka, *J. Magn. Magn. Mater.* 139, L231 (1995).
29. J. S. Moodera, L. R. Kinder, T. M. Wong, and R. Meservey, *Phys. Rev. Lett.* 74, 3273 (1995); J. S. Moodera and L. R. Kinder, *J. Appl. Phys.* 79, 4724 (1996).
30. H. Ohno, *Science* 281, 951 (1998).
31. H. Ohno, F. Matsukura, and Y. Ohno, *JSAP Int.* 5, 4 (2002).
32. Diluted magnetic semiconductors, in "Semiconductors and Semimetals" (J. K. Furydna and J. Kossut, Eds.), Vol. 25. Academic, New York, 1988.
33. T. Dietl, *Semicond. Sci. Technol.* 17, 377 (2002).
34. S. Sato and H. Katayama-Yoshida, *Semicond. Sci. Technol.* 17, 367 (2002).
35. M. L. Reed, N. A. El-Masry, and H. H. Stadelmaier, *Appl. Phys. Lett.* 79, 3473 (2001).
36. T. Sakaki, S. Sonoda, Y. Yamamoto, S. Shimizu, K. Suga, K. Kindo, and H. Hori, *Appl. Phys. Lett.* 91, 7911 (2002).
37. G. A. Medvedkin, T. Ishibashi, T. Nishi, K. Hayata, Y. Hasegawa, and K. Sato, *Jpn. J. Appl. Phys.* 39, L949 (2000).
38. N. Theodoropoulou, A. F. Hebard, M. E. Overberg, C. R. Abernathy, S. J. Pearton, S. N. G. Chu, and R. G. Wilson, *Phys. Rev. Lett.* 89, 107203 (2002).
39. Y. Matsumoto, M. Murakami, T. Shono, T. Hasegawa, T. Fukumura, M. Kawasaki, P. Ahmet, T. Chikyow, S. Koshihara, and H. Koinuma, *Science* 291, 854 (2001).
40. K. Ueda, H. Tabata, and T. Kawai, *Phys. Lett.* 79, 988 (2001).
41. M. Tanaka and Y. Higo, *Phys. Rev. Lett.* 87, 026602 (2001).
42. M. Tanaka, *Semicond. Sci. Technol.* 17, 327 (2002).
43. G. A. Prinz, *Phys. Today* 48, 58 (1995).
44. R. Fiederling, M. Keim, G. Reuscher, W. Ossau, G. Schmidt, A. Waag, and L. W. Molenkamp, *Nature* 402, 787 (1999).
45. T. Fujisawa, D. G. Austing, Y. Tokura, Y. Hirayama, and S. Tarucha, *Nature* 419, 278 (2002).
46. J. Lohau, A. Moser, C. T. Rettner, M. E. Best, and B. D. Terris, *IEEE Trans. Magn.* 37, 1652 (2001).
47. C. A. Ross, *Annu. Rev. Mater. Res.* 31, 203 (2001).
48. S. Sun, C. B. Murray, D. Weller, L. Folks, and A. Moser, *Science* 87, 1989 (2000).
49. International technology roadmap for semiconductors, 2001; http://public.itrs.net.
50. W. Thomson, *Proc. R. Soc.* 8, 546 (1857).
51. R. I. Potter, *Phys. Rev. B* 10, 4626 (1974).
52. T. R. McGuire and R. I. Potter, *IEEE Trans. Magn.* MAG-11, 1018 (1975).
53. R. Hunt, *IEEE Trans. Magn.* MAG-7, 150 (1971).
54. D. A. Thompson, *AIP Conf. Proc.* 24, 528 (1974).
55. J. C. Mallinson, "Magneto-Resistive Heads: Fundamentals and Applications." Academic, New York, 1996.
56. F. Jeffers, J. Freeman, R. Toussant, N. Smith, D. Wachenschwanz, S. Shtrikman, and W. Doyle, *IEEE Trans. Magn.* MAG-21, 1563 (1985).
57. K. Yamada, T. Maruyama, T. Tatsumi, T. Suzuki, K. Shimabayashi, Y. Motomura, M. Aoyama, and H. Urai, *IEEE Trans. Magn.* 26, 3010 (1990).
58. F. B. Shelledy, *IEEE Trans. Magn.* 28, 2283 (1992).
59. K. B. Klaassen and J. C. L. van Peppen, *IEEE Trans. Magn.* 33, 2611 (1997).
60. N. H. Yeh, *IEEE Trans. Magn.* MAG-18, 1155 (1982).
61. D. E. Heim, *IEEE Trans. Magn.* 30, 1453 (1994).
62. "Molecular Beam Epitaxy" (A. Cho, Ed.). AIP, New York, 1994.
63. L. L. Chang, L. Esaki, W. E. Howard, and R. Ludeke, *J. Vac. Sci. Technol.* 10, 11 (1973).
64. T. Shinjo and T. Takada, in "Ferromagnetic Materials" (E. P. Wohlfarth, Ed.), Vol. 3. Elsevier, Amsterdam, 1987.
65. N. F. Mott and H. Jones, "The Theory of the Properties of Metals and Alloys," pp. 265–268. Clarendon, Oxford, 1936.
66. M. A. M. Gijs and G. E. W. Bauer, *Adv. Phys.* 46, 285 (1997).
67. J. Mathon, in "Spin Electronics" (M. Ziese and M. J. Thornton, Eds.), Chap. 4, p. 75 and references therein. Springer, Berlin, 2001.
68. R. E. Camley and J. Barnas, *Phys. Rev. Lett.* 63, 664 (1989).
69. J. Barnas, A. Fuss, R. E. Camley, P. Grünberg, and W. Zinn, *Phys. Rev. B* 42, 8110 (1990).
70. R. Q. Hood and L. M. Falicov, *Phys. Rev. B* 46, 8287 (1992).
71. S. Zhang and P. M. Levy, *J. Appl. Phys.* 69, 4786 (1991).
72. P. M. Levy, *Solid State Phys.* 47, 367 (1994).
73. P. M. Levy and S. Zhang, *J. Magn. Magn. Mater.* 151, 315 (1995).
74. X. G. Zhang and W. H. Butler, *Phys. Rev. B* 51, 10085 (1995).
75. P. Zahn, I. Mertig, M. Richter, and H. Eschrig, *Phys. Rev. Lett.* 75, 2996 (1995).
76. M. A. M. Gijs and G. E. W. Bauer, *Adv. Phys.* 46, 285 (1997).
77. R. K. Nesbet, *IBM J. Res. Develop.* 42, 53 (1998).
78. P. Grünberg, R. Schreiber, Y. Pang, M. B. Brodsky, and H. Sowers, *Phys. Rev. Lett.* 57, 2442 (1986).
79. S. S. P. Parkin, N. More, and K. P. Roche, *Phys. Rev. Lett.* 64, 2304 (1990).
80. S. S. P. Parkin, R. Bhadra, and K. P. Roche, *Phys. Rev. Lett.* 66, 2152 (1991).
81. S. S. P. Parkin, *Phys. Rev. Lett.* 67, 3598 (1991).
82. S. S. P. Parkin, in "Ultrathin Magnetic Structures II" (B. Heinrich and J. A. C. Bland, Eds.), pp. 148–186. Springer, Berlin, 1994.
83. H. Hosoito, K. Mibu, S. Araki, T. Shinjo, S. Itoh, and Y. Endo, *J. Phys. Soc. Jpn.* 61, 300 (1992).
84. P. Grunberg, S. Demokritov, A. Fuss, R. Schreiber, J. A. Wolf, and S. T. Purcell, *J. Magn. Magn. Mater.* 104–107, 1734 (1992).
85. A. Fuss, S. Demokritov, P. Grünberg, and W. Zinn, *J. Magn. Magn. Mater.* 103, L221 (1992).
86. J. Unguris, R. J. Celotta, and D. T. Pierce, *Phys. Rev. Lett.* 67, 140 (1991).
87. J. A. Wolf, Q. Leng, R. Schreiber, P. A. Grünberg, and W. Zinn, *J. Magn. Magn. Mater.* 121, 253 (1993).
88. A. Cebollada, R. Miranda, C. M. Schneider, P. Schuster, and J. Kirschner, *J. Magn. Magn. Mater.* 102, 25 (1991).
89. Z. Q. Qiu, J. Pearson, and S. D. Bader, *Phys. Rev. B* 46, 8659 (1992).
90. M. T. Johnson, S. T. Purcell, N. W. E. McGee, R. Coehoorn, J. aan de Stegge, and W. Hoving, *Phys. Rev. Lett.* 68, 2688 (1992).
91. D. M. Edwards, J. Mathon, R. B. Muniz, and M. S. Phan, *Phys. Rev. Lett.* 67, 493 (1991).
92. P. Lang, L. Nordström, R. Zeller, and P. H. Dederichs, *Phys. Rev. Lett.* 71, 1927 (1993).
93. L. Nordström, P. Lang, R. Zeller, and P. H. Dederichs, *Phys. Rev. B* 50, 13058 (1994).

94. B. A. Jones and C. B. Hanna, *Phys. Rev. Lett.* 71, 4253 (1993).
95. J. Mathon, Murielle Villeret, R. B. Muniz, and J. d'Albuquerque e Castro, *Phys. Rev. Lett.* 74, 3696 (1995).
96. M. D. Stiles, *Phys. Rev. B* 48, 7238 (1993).
97. P. Bruno, *Phys. Rev. B* 52, 411 (1995).
98. Y. Yafet, *Phys. Rev. B* 36, 3948 (1987).
99. R. Coehoorn, *Phys. Rev. B* 44, 9331 (1991).
100. P. Bruno and C. Chappert, *Phys. Rev. Lett.* 67, 1602 (1991); P. Bruno and C. Chappert, *Phys. Rev. Lett.* 67, 2592 (1991).
101. P. Bruno and C. Chappert, *Phys. Rev. B* 46, 261 (1992)
102. A. Fert and P. Bruno, in "Ultrathin Magnetic Structures II" (B. Heinrich and J. A. C. Bland, Eds.), pp. 82–117. Springer, Berlin, 1994.
103. C. Chappert and J. P. Renard, *Europhys. Lett.* 15, 553 (1991).
104. D. M. Deaven, D. S. Rokhsar, and M. Johnson, *Phys. Rev. B* 44, 5977 (1991).
105. T. Shinjo and H. Yamamoto, *J. Phys. Soc. Jpn.* 59, 3061 (1990).
106. Y. Kawawake, H. Sakakima, Y. Irie, and M. Satomi, *J. Magn. Magn. Mater.* 156, 405 (1996).
107. S. S. P. Parkin, *Phys. Rev. Lett.* 71, 1641 (1993).
108. H. Yamamoto, Y. Motomura, T. Anno, and T. Shinjo, *J. Magn. Magn. Mater.* 126, 437 (1993).
109. A. V. Pohm, B. A. Everitt, R. S. Beech, and J. M. Daugton, *IEEE Trans. Magn.* 33, 3280 (1997).
110. A. A. Seigler, P. E. Anderson, and A. M. Shukh, *J. Appl. Phys.* 91, 2176 (2002).
111. Y. H. Wu, K. B. Li, J. J. Qiu, Z. B. Guo, and G. C. Han, *Appl. Phys. Lett.* 80, 4413 (2002).
112. B. Dienyl, V. S. Speriosu, B. A. Gurney, S. S. P. Parkin, D. R. Wilhoit, K. P. Roche, S. Metin, D. T. Peterson, and S. Nadimi, *J. Magn. Magn. Mater.* 93, 101 (1991).
113. B. Dieny, V. S. Speriosu, S. Metin, S. S. P. Parkin, B. A. Gurney, P. Baumgart, and D. R. Wilhoit, *J. Appl. Phys.* 69, 4774 (1991).
114. H. Kanai, Y. Yamada, K. Aoshima, Y. Ohtsuka, J. Kane, M. Kanamine, J. Toda, and Y. Mizoshita, *IEEE Trans. Magn.* 32, 3368 (1996).
115. S. S. P. Parkin, *Appl. Phys. Lett.* 61, 1358 (1992).
116. C. Tsang, R. E. Fontana, T. Lin, D. E. Heim, V. S. Speriosu, B. A. Gurney, and M. L. Williams, *IEEE Trans. Magn.* 30, 3801 (1994).
117. M. A. M. Gijs, J. B. Giesbers, and S. K. J. Lenczowski, *Phys. Rev. Lett.* 70, 3343 (1993).
118. W. Vavra, S. F. Cheng, A. Fink, J. J. Krebs, and G. A. Prinz, *Appl. Phys. Lett.* 66, 2579 (1995).
119. A. Blondel, J. P. Meier, B. Doudin, and J.-Ph. Ansermet, *Appl. Phys. Lett.* 65, 3019 (1994).
120. L. Piraux, J. M. George, J. F. Despres, C. Leroy, E. Ferain, R. Legras, K. Ounadjela, and A. Fert, *Appl. Phys. Lett.* 65, 2488 (1994).
121. K. Liu, K. Nagodawithana, P. C. Searson, and C. L. Chien, *Phys. Rev.* B51, 738 (1985).
122. T. Ono and T. Shinjo, *J. Phys. Soc. Jpn.* 64, 363 (1995).
123. E. M. Williams, "Design and Analysis of Magnetoresistive Recording Heads," p. 161. Wiley, New york, 2001.
124. S. Yuan and H. N. Bertram, *J. Appl. Phys.* 75, 6385 (1994).
125. D. E. Heim, R. E. Fontana, C. Tsang, V. S. Speriosu, B. A. Gurney, and M. L. Williams, *IEEE Trans. Magn.* 30, 316 (1994).
126. W. P. Meiklejohn and C. P Bean, *Phys. Rev.* 102, 1413 (1956).
127. J. Nogués and I. K. Schuller, *J. Magn. Magn. Mater.* 192, 203 (1999) and references therein.
128. A. E. Berkowitz and K. Takano, *J. Magn. Magn. Mater.* 200, 552 (1999) and references therein.
129. R. L. Stamps, *J. Phys. D: Appl. Phys.* 33, R247 (2000) and references therein.
130. M. Kiwi, *J. Magn. Magn. Mater.* 234, 584 (2001) and references therein.
131. W. P. Meiklejohn, *J. Appl. Phys.* 33, 1328 (1962).
132. L. Néel, *Ann. Phys. Paris* 2, 61 (1967).
133. A. P. Malozemoff, *Phys. Rev. B* 35, 3679 (1987).
134. D. Mauri, H. C. Siegmann, P. S. Bagus, and E. Kay, *J. Appl. Phys.* 62, 3047 (1987).
135. N. C. Koon, *Phys. Rev. Lett.* 78, 4865 (1997).
136. T. C. Schulthess and W. H. Butler, *Phys. Rev. Lett.* 81, 4516 (1998).
137. T. C. Schulthess and W. H. Butler, *J. Appl. Phys.* 85, 5510 (1999).
138. M. Kiwi, J. Mejía-López, R. D. Portugal, and R. Ramírez, *Appl. Phys. Lett.* 75, 3975 (1999).
139. S. Zhang, D. Dimitrov, G. C. Hadjipanayis, J. W. Cai, and C. L. Chien, *J. Magn. Magn. Mater.* 198–199, 468 (1999).
140. M. D. Stiles and R. D. McMicheal, *Phys. Rev. B* 59, 3722 (1999).
141. F. Y. Yang and C. L. Chien, *Phys. Rev. Lett.* 85, 2597 (2000).
142. P. Steadman, M. Ali, A. T. Hindmarch, C. H. Marrows, B. J. Hickey, S. Langridge, R. M. Dalgliesh, and S. Foster, *Phys. Rev. Lett.* 89, 077201 (2002).
143. H. Ohldag, T. J. Regan, J. Stöhr, A. Scholl, F. Nolting, J. Lüning, C. Stamm, S. Anders, and R. L. White, *Phys. Rev. Lett.* 87, 247201 (2001).
144. M. Jimbo, *J. Magn. Soc. Jpn.* 22, 12 (1998) (in Japanese).
145. M. Lederman, *IEEE Trans. Magn.* 35, 794 (1999).
146. T. Lin, C. Tsang, R. E. Fontana, and K. K. Howard, *IEEE Trans. Magn.* 31, 2585 (1995). The thickness of NiFe is 5 nm.
147. T. Lin, D. Mauri, N. Staud, C. Hwang, J. K. Howard, and G. Gorman, *Appl. Phys. Lett.* 65, 1183 (1994).
148. A. J. Devasahayam and M. H. Kryder, *IEEE Trans. Magn.* 35, 649 (1999). The thickness of NiFe is 25 nm.
149. G. W. Anderson, Y. Huai, and M. Pakala, *J. Appl. Phys.* 87, 5726 (2000). The free layer consists of 5 nm NiFe and 2 nm CoFe.
150. Z. Qian, J. M. Sivertsen, J. H. Judy, B. A. Everitt, S. Mao, and E. S. Murdock, *J. Appl. Phys.* 85, 6106 (1999).
151. S. Soeya, S. Tadokoro, T. Imagawa, M. Fuyama, and S. Narishige, *J. Appl. Phys.* 74, 6297 (1993).
152. M. J. Carey and A. E. Berkowitz, *Appl. Phys. Lett.* 65, 1183 (1994).
153. J. P. Nozières, S. Jaren, Y. B. Zhang, A. Zeltser, K. Pentek, and V. S. Speriosu, *J. Appl. Phys.* 87, 3920 (2000). The free layer is CoFe.
154. A. Veloso, P. P. Freitas, N. J. Oliveira, J. Fernandes, and M. Ferreira, *IEEE Trans. Magn.* 34, 2343 (1998).
155. M. Saito, Y. Kakihara, T. Watanabe, and N. Hasegawa, 21, 505 (1996) (in Japanese).
156. M. Sano, S. Araki, M. Ohta, K. Noguchi, H. Morita, and M. Matsuzaki, *IEEE Trans. Magn.* 34, 372 (1998).
157. S. Soeya, H. Hoshiya, M. Fuyama, and S. Tadokoro, *J. Appl. Phys.* 80, 1006 (1996).
158. M. Saito, Y. Kakihara, T. Watanabe, and N. Hasegawa, 21, 505 (1996) (in Japanese).
159. M. Sano, S. Araki, M. Ohta, K. Noguchi, H. Morita, and M. Matsuzaki, *IEEE Trans. Magn.* 34, 372 (1998).
160. S. Bae, J. H. Judy, P. J. Chen, W. F. Egelhoff, Jr., and S. Zurn, *Appl. Phys. Lett.* 78, 4163 (2001).
161. D. E. Heim and S. S. P. Parkin, U. S. Patent 5, 464, 185.
162. V. S. Sperious, B. A. Gurney, D. R. Wilhoit, and L. B. Brown, "Digest of INTERMAG '96," 1996, paper AA-04.
163. H. A. M. van den Berg, W. Clemens, G. Gieres, G. Rupp, W. Schelter, and M. Vieth, *IEEE Trans. Magn.* 32, 4624 (1996).
164. J. G. Zhu and Y. F. Zheng, *IEEE Trans. Magn.* 34, 1063 (1998).
165. J. L. Leal and M. H. Kryder, *J. Appl. Phys.* 83, 794 (1999).
166. C. Tsang, M. Pinarbasi, H. Santini, E. Marinero, P. Arnett, R. Olson, R. Hsiao, M. Williams, R. Payne, R. Wang, J. Moore, B. Gurney, T. Lin, and R. Fontana, *IEEE Trans. Magn.* 35, 689 (1999).
167. J. G. Zhu, *IEEE Trans. Magn.* 35, 655 (1999).
168. K. Meguro, H. Hoshiya, K. Watanabe, Y. Hamakawa, M. Fuyama, and H. Fukui, *IEEE Trans. Magn.* 35, 2925 (1999).
169. H. Kanai, M. Kanamine, A. Hashimoto, K. Aoshima, K. Noma, M. Yamagishi, H. Ueno, Y. Uehara, and Y. Uematsu, *IEEE Trans. Magn.* 35, 2580 (1999).

170. X. Yan, D. Rao, M. R. Gibbons, C. Qian, and H. C. Tong, *IEEE Trans. Magn.* 35, 2877 (1999).
171. S. Mao, A. Mack, E. Singleton, J. Chen, S. S. Xue, H. Wang, Z. Gao, J. Li, and E. Murdoc, *J. Appl. Phys.* 87, 5720 (2000).
172. A. Veloso, P. P. Freitas, and L. V. Melo, *IEEE Trans. Magn.* 35, 2568 (1999).
173. S. Araki, M. Sano, S. Li, Y. Tsuchiya, O. Redon, T. Sasaki, N. Ito, K. Terunuma, H. Morita, and M. Matsuzaki, *J. Appl. Phys.* 87, 5377 (2000).
174. H. M. A. van den Berg, G. Rupp, W. Clemens, G. Gieres, M. Vieth, J. Wecker, and S. Zoll, *IEEE Trans. Magn.* 34, 1336 (1998).
175. M. J. Carey, S. Maat, P. Rice, R. F. C. Farrow, R. F. Marks, A. Kellock, P. Nguyen, and B. A. Gurney, *Appl. Phys. Lett.* 81, 1044 (2002).
176. C. H. Bajorek, L. T. Romankiw, and D. A. Thompson, U.S. Patent 3,840,898, 1974.
177. R. D. Hempstead, S. Krongelb, and D. A. Thompson, *IEEE Trans. Magn.* MAG-14, 521 (1978).
178. C. Tsang and R. E. Fontana, Jr., *IEEE Trans. Magn.* MAG-18, 1149 (1982).
179. C. Tsang, *IEEE Trans. Magn.* 25, 3692 (1989).
180. M. T. Krounbi, W. J. Van Gestel, and P.-K. Wang, U.S. Patent 5, 018, 037, 1991.
181. D. Hannon, M. Krounbi, and J. Christner, *IEEE Trans. Magn.* 30, 298 (1994).
182. S. H. Liao, T. Torng, and T. Kobayashi, *IEEE Trans. Magn.* 30, 3855 (1994).
183. H. Kanai, J. Kane, K. Aoshima, M. Kanamine, and Y. Uehara, *IEEE Trans. Magn.* 31, 2612 (1995).
184. R. E. Fontana, Jr., S. A. MacDonald, C. Tsang, and T. Lin, *IEEE Trans. Magn.* 32, 3440 (1996).
185. M. Xiao, A. J. Devasahayam, and M. H. Kryder, *IEEE Trans. Magn.* 34, 1495 (1998).
186. J. G. Zhu and D. J. O'Connor, *IEEE Trans. Magn.* 32, 3401 (1996).
187. K. Takano, N. Yamanaka, and M. Matsuzaki, *IEEE Trans. Magn.* 34, 1516 (1998).
188. E. Champion and H. N. Bertram, *IEEE Trans. Magn.* 32, 13 (1996).
189. C. Mitsumata, K. Kikuchi, and T. Kobayashi, *IEEE Trans. Magn.* 34, 1453 (1998).
190. N. J. Zhu, G. Hiner, A. Rana, Y. M. Huai, and D. Seagle, *IEEE Trans. Magn.* 36, 2575 (2000).
191. A. Morinaga, C. Ishikawa, T. Ohtsu, N. Miyamoto, and S. Narishige, *IEEE Trans. Magn.* 38, 2262 (2002).
192. T. Suzuki, K. Ishihara, and H. Matsutera, *IEEE Trans. Magn.* 32, 3383 (1996).
193. Y. K. Zheng, Y. H. Wu, and T. C. Chong, *IEEE Trans. Magn.* 36, 3158 (2000).
194. C. Ishikawa, K. Suzuki, K. Yoshida, Y. Sugita, K. Shinagawa, Y. Nakatani, and N. Hayashi, *J. Appl. Phys.* 75, 1036 (1994).
195. T. Imamura, H. Kanai, and J. Toda, *IEEE Trans. Magn.* 35, 2559 (1999).
196. J. C. S. Kools, K. Rook, H. Hegde, S. B. Sant, J. Wong, W. Xiong, B. Druz, A. Lam, A. Devayasaham, and I. Wagner, *Thin Solid Films* 377–378, 705 (2000).
197. N. Sharma, C. Rea, and W. O'Kane, *IEEE Trans. Magn.* 36, 2496 (2000).
198. S. Mao, Z. Gao, H. W. Xi, P. Kolbo, M. Plumer, L. Wang, A. Goyal, I. Jin, J. Chen, C. H. Hou, R. M. White, and E. Murdock, *IEEE Trans. Magn.* 38, 26 (2002).
199. J. G. Zhu, Y. F. Zheng, and S. Liao, *IEEE Trans. Magn.* 37, 1723 (2001).
200. K. Nakamoto, Y. Kawato, M. Komuro, Y. Hamakawa, and T. Kawabe, "Digest of INTERMAG 2000," Toronto, Canada, Apr. 2000, paper FA-06.
201. J. Zhang, Y. Huai, and M. Lederman, *J. Appl. Phys.* 91, 7285 (2002).
202. Y. H. Wu, Y. K. Zheng, and T. C. Chong, "Digest of INTERMAG 2003," Boston, MA, Mar. 2003, paper DD06.
203. O. Redon, N. Kasahara, K. Shimazawa, S. Araki, H. Morita, and M. Matsuzaki, *J. Appl. Phys.* 87, 4689 (2000).
204. K. Sin, M. R. Gibbons, S. Funada, M. Mao, D. Rao, C. Chien, and H. C. Tong, *J. Appl. Phys.* 89, 7359 (2001).
205. C. Tsang, *J. Appl. Phys.* 55, 2226 (1984).
206. X. Shi, K. Ju, J. Hagen, C. L. Lin, C. C. Han, M. M. Chen, J. W. Chang, P. W. Wang, and E. Teng, *IEEE Trans. Magn.* 33, 2896 (1997).
207. C. Ishikawa, K. Suzuki, N. Koyama, K. Yoshida, Y. Sugita, K. Shinagawa, Y. Nakatani, and N. Hayashi, *J. Appl. Phys.* 74, 5666 (1993).
208. Y. Shen, W. Chen, W. Jensen, D. Ravipati, V. Retort, R. Rottmayer, S. Rudy, S. Tan, and S. Yuan, *IEEE Trans. Magn.* 32, 19 (1996).
209. A. M. Mack, K. Subramanian, L. R. Pust, C. J. Rea, N. Amin, M. A. Seigler, S. Mao, S. Xue, and S. Gangopadhyay, *IEEE Trans. Magn.* 37, 1727 (2001).
210. E. Nakashio, J. Sugawara, S. Onoe, and S. Kumagai, *J. Appl. Phys.* 89, 7356 (2001).
211. S. Mao, Z. Gao, H. W. Xi, P. Kolbo, M. Plumer, L. Wang, A. Goyal, I. Jin, J. Chen, C. H. Hou, R. M. White, and E. Murdock, *IEEE Trans. Magn.* 38, 26 (2002).
212. N. J. Gokemeijer, T. Ambrose, and C. L. Chien, *Phys. Rev. Lett.* 79, 4270 (1997).
213. N. J. Gokemeijer, T. Ambrose, C. L. Chien, N. Wang, and K. K. Fung, *J. Appl. Phys.* 81, 4999 (1997).
214. B. A. Gurney, V. S. Speriosu, J.-P. Nozieres, H. Lefakis, D. R. Wilhoit, and O. U. Need, *Phys. Rev. Lett.* 71, 4023 (1993).
215. Y. Kamiguchi, H. Yuasa, H. Fukuzawa, K. Koi, H. Iwasaki, and M. Sahashi, "Digest of INTERMAG'99," 1999, paper DB-01.
216. M. Ueno, H. Nishida, K. Mizukami, F. Hikami, K. Tabuchi, and T. Sawasaki, *IEEE Trans. Magn.* 36, 2572 (2000).
217. P. M. Baumgart, B. Dieny, B. A. Gurney, J. P. Nozieres, V. S. Speriosu, and D. R. Wilhoit, U.S. Patent 5,287,238, 1994.
218. T. C. Anthony, J. A. Brug, and S. Zhang, *IEEE Trans. Magn.* 30, 3819 (1994).
219. W. F. Egelhoff, Jr., T. Ha, R. D. K. Misra, Y. Kadmon, J. Nir, C. J. Powell, M. D. Stiles, R. D. McMichael, C.-L. Lin, J. M. Sivertsen, J. H. Judy, K. Takano, A. E. Berkowitz, T. C. Anthony, and J. A. Brug, *J. Appl. Phys.* 78, 273 (1995).
220. J.-G. Zhu, X.-G. Ye, S. W. Yuan, H.-C. Tong, and R. Rottmayer, *J. Appl. Phys.* 79, 5886 (1996).
221. H. C. Tong, X. Shi, F. Liu, C. Qian, Z. W. Dong, X. Yan, R. Barr, L. Miloslavsky, S. Zhou, J. Perlas, P. Prabhu, M. Mao, S. Funada, M. Gibbons, Q. Leng, J. G. Zhu, and S. Dey, *IEEE Trans. Magn.* 35, 2574 (1999).
222. A. Tanaka, Y. Shimizu, H. Kishi, K. Nagasaka, and M. Oshiki, *IEEE Trans. Magn.* 33, 3592 (1997).
223. A. Tanaka, Y. Shimizu, H. Kishi, K. Nagasaka, H. Kanai, and M. Oshiki, *IEEE Trans. Magn.* 35, 700 (1999).
224. B. Dieny, *Europhys. Lett.* 17, 261 (1992).
225. H. J. M. Swagten, G. J. Strijkers, P. J. H. Bloemen, M. M. H. Willekens, and W. J. M. de Jonge, *Phys. Rev. B* 53, 9108 (1996).
226. N. Hasegawa, A. Makino, F. Koike, and K. Ikarashi, *IEEE Trans. Magn.* 32, 4618 (1996).
227. Y. Sugita, Y. Kawawake, M. Satomi, and H. Sakakima, *Jpn. J. Appl. Phys.* 37, 5984 (1998).
228. Y. Kawawake, Y. Sugita, M. Sotomi, and H. Sakakima, *J. Appl. Phys.* 85, 5024 (1999).
229. S. Bae, J. H. Judy, W. F. Egelhoff, Jr., and P. J. Chen, *J. Appl. Phys.* 87, 6980 (2000).
230. W. F. Egelhoff, Jr., P. J. Chen, C. J. Powell, M. D. Stiles, R. D. McMichael, J. H. Judy, K. Takano, A. E. Berkowitz, and J. M. Daughton, *IEEE Trans. Magn.* 33, 3580 (1997).

231. Y. Kawawake and H. Sakakima, *IEEE Trans. Magn.* 33, 3538 (1997).
232. S. Maat, M. J. Carey, E. E. Fullerton, T. X. Le, P. M. Rice, and B. A. Gurney, *Appl. Phys. Lett.* 81, 520 (2002).
233. W. F. Egelhoff, Jr., P. J. Chen, C. J. Powell, M. D. Stiles, R. D. McMichael, J. H. Judy, K. Takano, and A. E. Berkowitz, *J. Appl. Phys.* 82, 6142 (1997).
234. H. Sakakima, M. Satomi, Y. Sugita, and Y. Kawawake, *J. Magn. Magn. Mater.* 210, 20 (2000).
235. T. Kato, K. Miyashita, S. Iwata, S. Tsunashima, H. Sakakima, Y. Sugita, and Y. Kawawake, *J. Magn. Magn. Mater.* 240, 168 (2002).
236. J. Hong, K. Aoshima, J. Kane, K. Noma, and H. Kanai, *IEEE Trans. Magn.* 36, 2629 (2000).
237. Y. Tsuchiya, S. Li, M. Sano, T. Uesugi, S. Araki, H. Morita, and M. Matsuzaki, *IEEE Trans. Magn.* 36, 2557 (2000).
238. M. Mao, C. Cerjan, B. Law, F. Grabner, L. Miloslavsky, and C. Chien, *IEEE Trans. Magn.* 36, 2866 (2000).
239. M. F. Gillies and A. E. T. Kuiper, *J. Appl. Phys.* 88, 5894 (2000).
240. A. Veloso, P. P. Freitas, P. Wei, N. P. Barradas, J. C. Soares, B. Almeida, and J. B. Sousa, *Appl. Phys. Lett.* 77, 1020 (2000).
241. H. Kanai, K. Noma, and J. Hong, *Fujitsu Sci. Tech. J.* 37, 174 (2001).
242. J. Hong, K. Noma, and H. Kanai, *J. Appl. Phys.* 89, 6940 (2001).
243. H. Fukuzawa, K. Koi, H. Tomita, H. N. Fuke, Y. Kamiguchi, H. Iwasaki, and M. Sahashi, *J. Magn. Magn. Mater.* 235, 208 (2001).
244. T. Lin and D. Mauri, *Appl. Phys. Lett.* 78, 2181 (2001).
245. S. Sant, M. Mao, J. Kools, K. Koi, H. Iwasaki, and M. Sahashi, *J. Appl. Phys.* 89, 6931 (2001).
246. J. C. S. Kools, S. B. Sant, K. Book, W. Xiong, Faiz Dahmani, W. Ye, J. Nuñez-Regueiro, Y. Kawana, M. Mao, K. Koi, H. Iwasaki, and M. Sahashi, *IEEE Trans. Magn.* 37, 1783 (2001).
247. Y. Sugita, Y. Kawawake, M. Satomi, and H. Sakakima, *J. Appl. Phys.* 89, 6919 (2001).
248. M. F. Gillies, A. E. T. Kuiper, and G. W. R. Leibbrandt, *J. Appl. Phys.* 89, 6922 (2001).
249. K. Li, Y. Wu, J. Qiu, G. Han, Z. Guo, H. Xie, and T. Chong, *Appl. Phys. Lett.* 79, 3663 (2001).
250. K. Li, Y. Wu, J. Qiu, and T. Chong, *J. Appl. Phys.* 91, 8563 (2002).
251. J. B. Sousa, J. O. Ventura, M. A. Salgueiro da Silva, P. P. Freitas, and A. Veloso, *J. Appl. Phys.* 91, 5321 (2002).
252. S. H. Jang, T. Kang, H. J. Kim, and K. Y. Kim, *J. Magn. Magn. Mater.* 240, 192 (2002).
253. S. H. Jang, T. Kang, H. J. Kim, and K. Y. Kim, *Appl. Phys. Lett.* 81, 105 (2002).
254. A. Al-Jibouri, M. Hoban, Z. Lu, and G. Pan, *J. Appl. Phys.* 91, 7098 (2002).
255. F. Shen, Q. Y. Xu, G. H. Yu, W. Y. Lai, Z. Zhang, Z. Q. Lu, G. Pan, and A. Al-Jibouri, *Appl. Phys. Lett.* 80, 4410 (2002).
256. M. Mao, C. Cerjan, and J. Kools, *J. Appl. Phys.* 91, 8560 (2002).
257. Y. Huai, Z. T. Diao, and J. Zhang, *IEEE Trans. Magn.* 38, 20 (2002).
258. N. Hasegawa, F. Koike, K. Ikarashi, M. Ishizone, M. Kawamura, Y. Nakazawa, A. Takahashi, H. Tomita, H. Iwasaki, and M. Sahashi, *J. Appl. Phys.* 91, 8774 (2002).
259. H. Fukuzawa, K. Koi, H. Tomita, H. N. Fuke, H. Iwasaki, and M. Sahashi, *J. Appl. Phys.* 91, 6684 (2002).
260. J. Hong, J. Kane, J. Hashimoto, M. Yamagishi, K. Noma, and H. Kanai, *IEEE Trans. Magn.* 38, 15 (2002).
261. T. Mizuguchi and H. Kano, *IEEE Trans. Magn.* 37, 1742 (2001).
262. W. Y. Lee, M. Carey, M. F. Toney, P. Rice, B. Gurney, H.-C. Chang, E. Allen, and D. Mauri, *J. Appl. Phys.* 89, 6925 (2001).
263. W. C. Uhlig, M. Mao, V. Yiu, J. Li, and J. Shi, *J. Appl. Phys.* 89, 6937 (2001).
264. C.-H. Lai, C. J. Chen, and T. S. Chin, *J. Appl. Phys.* 89, 6928 (2001).
265. L. Néel, *C. R. Acad. Sci.* 255, 1676 (1962).
266. J. C. S. Kools, W. Kula, D. Mauri, and T. Lin, *J. Appl. Phys.* 85, 4466 (1999).
267. S. Miura, M. Tsunoda, and M. Takahashi, *J. Appl. Phys.* 89, 6308 (2001).
268. R. Wood, *IEEE Trans. Magn.* 36, 36 (2000).
269. R. Rottmayer and J. G. Zhu, *IEEE Trans. Magn.* 31, 2597 (1995).
270. K. Nagasaka, Y. Seyama, L. Varga, Y. Shimizu, and A. Tanaka, *J. Appl. Phys.* 89, 6943 (2001).
271. M. Takahashi, T. Funayama, K. Tateyama, M. Yoshikawa, H. Iwasaki, and M. Sahashi, "Digest of MMM 2001," 2001, paper CB-01.
272. H. Yuasa, M. Yoshikawa, Y. Kamiguchi, K. Koi, H. Iwasaki, M. Takagishi, and M. Sahashi, *J. Appl. Phys.* 92, 2646 (2002).
273. A. Tanaka, Y. Shimizu, Y. Seyama, K. Nagasaka, R. Kondo, H. Oshima, S. Eguchi, and R. Kanai, *IEEE Trans. Magn.* 38, 84 (2002).
274. K. Nagasaka, Y. Seyama, R. Kondo, H. Oshima, Y. Shimizu, and A. Tanaka, *Fujitsu Sci. Tech. J.* 37, 192 (2001).
275. M. Takahashi, K. Koi, M. Yoshikawa, T. Funayama, H. Iwasaki, and M. Sahashi, "Digest of INTERMAG Europe 2002," Amsterdam, The Netherlands, Apr. 2002, Paper CA02.
276. K. Nagasaka, H. Oshima, Y. Seyama, Y. Shimizu, and A. Tanaka, "Digest of INTERMAG Europe 2002," Amsterdam, The Netherlands, Apr. 2002, paper AC04.
277. Y. H. Wu, K. B. Li, Z. B. Guo, J. J. Qiu, G. C. Han, and T. C. Chong, "Digest of INTERMAG Europe 2002," Amsterdam, The Netherlands, Apr. 2002, paper CA04.
278. A. Matsuzono, S. Terada, H. Ono, A. Furukawa, T. Sone, S. Sasaki, Y. Kakihara, Y. Takeda, N. Chiyokubo, and H. Matsuki, *J. Appl. Phys.* 91, 7267 (2002).
279. M. Hosomi, E. Makino, I. Konishiike, N. Sugawara, and S. Ohkawara, *J. Appl. Phys.* 91, 8099 (2002).
280. W. P. Pratt, Jr., S.-F. Lee, J. M. Slaughter, R. Loloee, P. A. Schroeder, and J. Bass, *Phys. Rev. B* 66, 3060 (1991).
281. S. F. Lee, W. P. Pratt, Jr., R. Loloee, P. A. Schroeder, and J. Bass, *Phys. Rev. B* 46, 548 (1992).
282. G. E. W. Bauer, *Phys. Rev. Lett.* 69, 1676 (1992).
283. S. Zhang and P. M. Levy, *J. Appl. Phys.* 69, 4786 (1991).
284. T. Valet and A. Fert, *Phys. Rev. B* 48, 274 (1999).
285. S.-F. Lee, W. P. Pratt, Jr., Q. Yang, P. Holody, R. Loloee, P. A. Schroeder, and J. Bass, *J. Magn. Magn. Mater.* 118, L1 (1993).
286. M. A. M. Gijs, S. K. J. Lenczowski, and J. B. Giesbers, *Phys. Rev. Lett.* 70, 3343 (1993).
287. M. A. M. Gijs, M. T. Johnson, A. Reinders, P. E. Huisman, R. J. M. van de Veerdonk, S. K. J. Lenczowski, and R. M. J. van Gansewinkel, *Appl. Phys. Lett.* 66, 1839 (1995).
288. P. M. Levy, S. Zhang, T. Ono, and T. Shinjo, *Phys. Rev. B* 52, 16049 (1995).
289. A. Vedyayev, M. Chshiev, N. Ryzhanova, B. Dieny, C. Cowache, and F. Brouers, *J. Magn. Magn. Mater.* 172, 53 (1997).
290. A. C. Reilly, W. Park, R. Slater, B. Ouaglal, R. Loloee, W. P. Pratt, Jr., and J. Bass, *J. Magn. Magn. Mater.* 195, L269 (1999).
291. J. Bass and W. P. Pratt, Jr., *J. Magn. Magn. Mater.* 200, 274 (1999).
292. J. Y. Gu, S. D. Steenwyk, A. C. Reilly, W. Park, R. Loloee, J. Bass, and W. P. Pratt, Jr., *J. Appl. Phys.* 87, 4831 (2000).
293. T. Takagi, F. Koyama, and K. Iga, *Appl. Phys. Lett.* 59, 2877 (1991).
294. C. Stirtori, F. Capasso, J. Faist, D. L. Sivco, S.-N. G. Chu, and A. Y. Cho, *Appl. Phys. Lett.* 61, 898 (1992).
295. F. J. Himpsel, J. E. Ortega, G. J. Mankey, and R. F. Willis, *Adv. Phys.* 47, 511 (1998) and references therein.
296. S. Yuasa, T. Nagahama, and Y. Suzuki, *Science* 297, 234 (2002).
297. R. Meservey and P. M. Tedrow, *Phys. Rep.* 238, 173 (1994).
298. J. S. Moodera, J. Nassar, and G. Mathon, *Annu. Rev. Mater. Sci.* 29, 381 (1999).
299. W. J. Gallagher, S. S. P. Parkin, Y. Lu, X. P. Bian, A. Marley, K. P. Roche, R. A. Altman, S. A. Rishton, C. Jahnes, T. M. Shaw, and G. Xiao, *J. Appl. Phys.* 81, 3741 (1997).

300. P. M. Tedrow and R. Moservey, *Phys. Rev. Lett.* 26, 192 (1971).
301. R. J. Soulen, Jr., J. M. Byers, M. S. Osofsky, B. Nadgorny, T. Ambrose, S. F. Cheng, P. R. Broussard, C. T. Tanaka, J. Nowak, J. S. Moodera, A. Barry, and J. M. D. Coey, *Science* 282, 85 (1998).
302. S. K. Upadhyay, A. Palanisami, R. N. Louie, and R. A. Buhrman, *Phys. Rev. Lett.* 81, 3247 (1998).
303. M. Bowen, M. Bibes, A. Barthélémy, J.-P. Contour, A. Anane, Y. Lemaître, and A. Fert, *Appl. Phys. Lett.* 82, 233 (2003).
304. G. Hu and Y. Suzuki, *Phys. Rev. Lett.* 89, 276601 (2002).
305. Rager, A. V. Berenov, L. F. Cohen, W. R. Branford, Y. V. Bugoslavsky, Y. Miyoshi, M. Ardakani, and J. L. MacManus-Driscoll, *Appl. Phys. Lett.* 81, 5003 (2002).
306. M. Sharma, S. X. Wang, and J. H. Nickel, *Phys. Rev. Lett.* 82, 616 (1999).
307. J. M. De Teresa, A. Barthélémy, A. Fert, J. P. Contour, R. Lyonnet, F. Montaigne, P. Seneor, and A. Vaurès, *Phys. Rev. Lett.* 82, 4288 (1999).
308. J. H. Park, E. Vescovo, H.-J. Kim, C. Kwon, R. Ramesh, and T. Venkatesan, *Nature* (*London*) 392, 794 (1998).
309. A. Ferta, A. Barthélémya, J. Ben Youssefc, J.-P. Contoura, V. Crosa, J. M. De Teresaa, A. Hamzica, J. M. Georgea, G. Faini, b, J. Grolliera, H. Jaffrèsa, H. Le Gallc, F. Montaignea, F. Paillouxa, and F. Petroffa, *Mater. Sci. Eng. B—Solid* 84, 1 (2001).
310. K. Ghosh, S. B. Ogale, S. P. Pai, M. Robson, E. Li, I. Jin, Z. Dong, R. L. Greene, R. Ramesh, T. Venkatesan, and M. Johnson, *Appl. Phys. Lett.* 73, 689 (1998).
311. A. Gupta, X. W. Li, and G. Xiao, *Appl. Phys. Lett.* 78, 1894 (2001).
312. P. K. de Boer, G. A. de Wijs, and R. A. de Groot, *Phys. Rev. B* 58, 15422 (1998).
313. I. I. Oleinik, E. Y. Tsymbal, and D. G. Pettifor, *Phys. Rev. B* 62, 3952 (2000).
314. E. Y. Tsymbal, I. I. Oleinik, and D. G. Pettifor, *J. Appl. Phys.* 87, 5230 (2000).
315. B. D. Schrag, A. Anguelouch, S. Ingvarsson, G. Xiao, Y. Lu, P. L. Trouilloud, A. Gupta, R. A. Wanner, W. J. Gallagher, P. M. Rice, and S. S. P. Parkin, *Appl. Phys. Lett.* 77, 2373 (2000).
316. L. F. Li, X. Y. Liu, and G. Xiao, *J. Appl. Phys.* 93, 467 (2003).
317. S. Tegen, I. Monch, J. Schumann, H. Vinzelberg, and C. M. Schneider, *J. Appl. Phys.* 89, 8169 (2001).
318. C. Tiusan, M. Hehn, and K. Ounadjela, *Eur. Phys. J. B* 26, 431 (2002).
319. J. J. Sun, K. Shimazawa, N. Kasahara, K. Sato, T. Kagami, S. Saruki, S. Araki, and M. Matsuzaki, *J. Appl. Phys.* 89, 6653 (2001).
320. J. E. Ortega, A. Närmann, K. N. Altmann, W. O'Brien, D. J. Seo, F. J. Himpsel, P. Segovia, A. Mascaraque, and E. G. Michel, *J. Magn. Magn. Mater.* 203, 126 (1999).
321. R. K. Kawakami, E. Rotenberg, E. J. Escorcia-Aparicio, H. J. Choi, T. R. Cummins, J. G. Tobin, N. V. Smith, and Z. Q. Qiu, *Phys. Rev. Lett.* 80, 1754 (1998).
322. A. Vedyayev, N. Ryzhanova, C. Lacroix, L. Giacomoni, B. Dieny, *Europhys. Lett.* 39, 219 (1997).
323. J. Mathon and A. Umerski, *Phys. Rev. B* 60, 1117 (1999).
324. S. Zhang and P. M. Levy, *Phys. Rev. Lett.* 81, 5660 (1998).
325. J. S. Moodera, J. Nowak, L. R. Kinder, P. M. Tedrow, R. J. M. van de Veerdonk, B. A. Smits, M. van Kampen, H. J. M. Swagten, and W. J. M. de Jonge, *Phys. Rev. Lett.* 83, 3029 (1999).
326. S. S. P. Parkin, U.S. Patent 5,764,567, 1998.
327. J. J. Sun and P. P. Freitas, *J. Appl. Phys.* 85, 5264 (1999).
328. P. LeClair, H. J. M. Swagten, J. T. Kohlhepp, R. J. M. van de Veedonk, and W. J. M. de Jonge, *Phys. Rev. Lett.* 84, 2933 (2000).
329. J. C. Slonczewski, *Phys. Rev. B* 39, 6995 (1989).
330. P. LeClair, B. Hoex, H. Wieldraaijer, J. T. Kohlhepp, H. J. M. Swagten, and W. J. M. de Jonge, *Phys. Rev. B* 64, 100406 (2001).
331. R. Schad, D. Allen, G. Zangari, I. Zana, D. Yang, M. Tondra, and D. X. Wang, *Appl. Phys. Lett.* 76, 607 (2000).
332. M. F. Gillies, A. E. T. Kuiper, J. B. A. van Zon, and J. M. Sturm, *Appl. Phys. Lett.* 78, 3496 (2001).
333. P. Rottlander, M. Hehn, O. Lenoble, and A. Schuhl, *Appl. Phys. Lett.* 78, 3274 (2001).
334. L. S. Dorneles, R. L. Sommer, and L. F. Schelp, *J. Appl. Phys.* 91, 7971 (2002).
335. M. F. Gillies, A. E. T. Kuiper, J. B. A. van Zon, and J. M. Sturm, *Appl. Phys. Lett.* 78, 3496 (2001).
336. M. M. Schwickert, J. R. Childress, R. E. Fontana, A. J. Kellock, P. M. Rice, M. K. Ho, T. J. Thompson, and B. A. Gurney, *J. Appl. Phys.* 89, 6871 (2001).
337. J. Wang, S. Cardoso, P. P. Freitas, P. Wei, N. P. Barradas, and J. C. Soares, *J. Appl. Phys.* 89, 6868 (2001).
338. P. Shang, A. K. Petford-Long, J. H. Nickel, M. Sharma, and T. C. Anthony, *J. Appl. Phys.* 89, 6874 (2001).
339. H. J. Shim, I. J. Hwang, K. S. Kim, B. K. Cho, J. T. Kim, and J. H. Sok, *J. Appl. Phys.* 92, 1095 (2002).
340. J. Wang, P. P. Freitas, E. Snoeck, X. Battle, and J. Cuadra, *J. Appl. Phys.* 91, 7463 (2002).
341. J. Wang, P. P. Freitas, E. Snoeck, P. Wei, and J. C. Soares, *Appl. Phys. Lett.* 79, 4387 (2001).
342. S. Kreuzer, W. Wegscheider, and D. Weiss, *J. Appl. Phys.* 89, 6751 (2001).
343. Z. S. Li, C. de Groot, and J. H. Moodera, *Appl. Phys. Lett.* 77, 3630 (2000).
344. F. J. Cadieu, L. Chen, and B. Li, *J. Appl. Phys.* 91, 7968 (2002).
345. M. Guth, V. Da Costa, G. Schmerber, A. Dinia, and H. A. M. van den Berg, *J. Appl. Phys.* 89, 6748 (2001).
346. M. Guth, G. Schmerber, Y. Henry, and A. Dinia, *J. Magn. Magn. Mater.* 240, 152 (2002).
347. Y. Lu, X. W. Li, G. Q. Gong, G. Xiao, A. Gupta, P. Lecoeur, J. Z. Sun, Y. Y. Wang, and V. P. Dravid, *Phys. Rev. B* 54, R8357–R8360 (1996).
348. T. Obata, T. Manako, Y. Shimakawa, and Y. Kubo, *Appl. Phys. Lett.* 74, 290 (1999).
349. J. Hayakawa, K. Ito, S. Kokado, M. Ichimura, A. Sakuma, M. Sugiyama, H. Asano, and M. Matsui, *J. Appl. Phys.* 91, 8792 (2002).
350. D. C. Worledge and T. H. Geballe, *Appl. Phys. Lett.* 76, 900 (2000).
351. I. I. Oleinik, E. Y. Tsymbal, and D. G. Pettifor, *Phys. Rev. B* 65, 020401 (2002).
352. S. Mitani, T. Moriyama, and K. Takanashik, *J. Appl. Phys.* 91, 7200 (2002).
353. F. Gustavsson, J. M. George, V. H. Etgens, and M. Eddrief, *Phys. Rev. B* 64, 184422 (2001).
354. D. M. Borsa, S. Grachev, and D. O. Boerma, *IEEE Trans. Magn.* 38, 2709 (2002).
355. D. J. Keavney, E. E. Fullerton, and S. D. Bader, *J. Appl. Phys.* 81, 795 (1997).
356. T. S. Plaskett, P. P. Freitas, N. P. Barradas, M. F. da Silva, and J. C. Soares, *J. Appl. Phys.* 76, 6104 (1994).
357. E. Popova, J. Faure-Vincent, C. Tiusan, C. Bellouard, H. Fischer, M. Hehn, F. Montaigne, M. Alnot, S. Andrieu, A. Schuhl, E. Snoeck, and V. da Costa, *Appl. Phys. Lett.* 81, 1035 (2002).
358. E. Popova, J. Faure-Vincent, C. Tiusan, C. Bellouard, H. Fischer, M. Hehn, F. Montaigne, M. Alnot, S. Andrieu, A. Schuhl, E. Snoeck, and V. da Costa, *Appl. Phys. Lett.* 81, 1035 (2002).
359. T. Kiyomura, Y. Maruo, and M. Gomi, *J. Appl. Phys.* 88, 4768 (2000).
360. C. L. Platt, B. Dieny, and A. E. Berkowitz, *Appl. Phys. Lett.* 69, 2291 (1996).
361. D. J. Smith, M. R. McCartney, C. L. Platt, and A. E. Berkowitz, *J. Appl. Phys.* 83, 5154 (1998).
362. C. L. Platt, B. Dieny, and A. E. Berkowitz, *J. Appl. Phys.* 81, 5523 (1997).
363. L. S. Dorneles, R. L. Sommer, and L. F. Schelp, *J. Appl. Phys.* 91, 7971 (2002).

364. K. Matsuda, A. Kamijo, T. Mitsuzuka, and H. Tsuge, *J. Appl. Phys.* 85, 5261 (1999).
365. H. Tsuge and T. Mitsuzuka, *Appl. Phys. Lett.* 71, 3296 (1997).
366. K. Matsuda, A. Kamijo, T. Mitsuzuka, and H. Tsuge, *J. Appl. Phys.* 85, 5261 (1999).
367. H. Boeve, J. De Boeck, and G. Borghs, *J. Appl. Phys.* 89, 482 (2001).
368. Z. G. Zhang, P. P. Freitas, A. R. Ramos, N. P. Barradas, and J. C. Soares, *J. Appl. Phys.* 91, 8786 (2002).
369. K.-S. Moon, Y. Chen, and Y. Huai, *J. Appl. Phys.* 91, 7965 (2002).
370. P. LeClair, J. T. Kohlhepp, A. A. Smits, H. J. M. Swagten, B. Koopmans, and W. J. M. de Jonge, *J. Appl. Phys.* 87, 6070 (2000).
371. S. Cardoso, V. Gehanno, R. Ferreira, and P. P. Freitas, *IEEE Trans. Magn.* 35, 2952 (1999).
372. B. F. P. Roos, P. A. Beck, S. O. Demokritov, B. Hillebrands, and D. Ozkaya, *J. Appl. Phys.* 89, 6656 (2001).
373. K. Nishikawa, M. Tsunoda, S. Ogata, and M. Takahashi, *IEEE Trans. Magn.* 38, 2718 (2002).
374. K. S. Yoon, J. H. Park, J. H. Choi, J. Y. Yang, C. H. Lee, C. O. Kim, J. P. Hong, and T. W. Kang, *Appl. Phys. Lett.* 79, 1160 (2001).
375. K. S. Yoon, J. H. Park, J. Y. Yang, C. O. Kim, and J. P. Hong, *J. Appl. Phys.* 91, 7953 (2002).
376. O. Song, Y. M. Lee, C. S. Yoon, and C. K. Kim, *J. Appl. Phys.* 93, 1146 (2003).
377. K. Nishikawa, M. Tsunoda, S. Ogata, and M. Takahashi, *IEEE Trans. Magn.* 38, 2718 (2002).
378. D. Song, J. Nowak, and M. Covington, *J. Appl. Phys.* 87, 5197 (2000).
379. U. May, K. Samm, H. Kittur, J. Hauch, R. Calarco, U. Rüdiger, and G. Güntherodt, *Appl. Phys. Lett.* 78, 2026 (2001).
380. P. Rottländer, H. Kohlstedt, P. Grünberg, and E. Girgis, *J. Appl. Phys.* 87, 6067 (2000).
381. H. Boeve, E. Girgis, J. Schelten, J. De Boeck, and G. Borghs, *Appl. Phys. Lett.* 76, 1048 (2000).
382. B. F. P. Roos, P. A. Beck, S. O. Demokritov, B. Hillebrands, and D. Ozkaya *J. Appl. Phys.* 89, 6656 (2001).
383. K. Shimazawa, N. Kasahara, J. J. Sun, S. Araki, H. Morita, and M. Matsuzaki, *J. Appl. Phys.* 87, 5194 (2000).
384. B. Park and T. D. Lee, *J. Magn. Magn. Mater.* 226, 926 (2001).
385. T. Schneider, H. Hegde, K. Sin, S. Funada, and S. Shi, *IEEE Trans. Magn.* 38, 2724 (2002).
386. M. Tsunoda, K. Nishikawa, S. Ogata, and M. Takahashi, *Appl. Phys. Lett.* 80, 3135 (2002).
387. X. F. Han, M. Oogane, H. Kubota, Y. Ando, and T. Miyazaki, *Appl. Phys. Lett.* 77, 283 (2000).
388. J. S. Bae, K. H. Shin, and H. M. Lee, *J. Appl. Phys.* 91, 7947 (2002).
389. Y. Ando, M. Hayashi, S. Iura, K. Yaoita, C. C. Yu, H. Kubota, and T. Miyazaki, *J. Phys. D: Appl. Phys.* 35, 2415 (2002).
390. T. Mitsuzuka, K. Matsuda, A. Kamijo, and H. Tsuge, *J. Appl. Phys.* 85, 5807 (1999).
391. Y. Li and S. X. Wang, *J. Appl. Phys.* 91, 7950 (2002).
392. E. Y. Chen, R. Whig, J. M. Slaughter, D. Cronk, J. Goggin, G. Steiner, and S. Tehrani, *J. Appl. Phys.* 87, 6061 (2000).
393. B. J. Jönsson-Åkerman, R. Escudero, C. Leighton, S. Kim, I. K. Schuller, and D. A. Rabson, *Appl. Phys. Lett.* 77, 1870 (2000).
394. D. Allen, R. Schad, G. Zangari, I. Zana, D. Yang, M. Tondra, and D. Wang, *J. Appl. Phys.* 87, 5188 (2000).
395. D. Allen, R. Schad, G. Zangari, I. Zana, M. Tondra, D. Wang, and D. Reed, *J. Appl. Phys.* 89, 6662 (2001).
396. T. Dimopoulos, V. Da Costa, C. Tiusan, K. Ounadjela, and H. A. M. van den Berg, *J. Appl. Phys.* 89, 7371 (2001).
397. Y. Ando, H. Kameda, H. Kubota, and T. Miyazaki, *J. Appl. Phys.* 87, 5206 (2000).
398. E. Z. Luo, S. K. Wong, A. B. Pakhomov, J. B. Xu, I. H. Wilson, and C. Y. Wong, *J. Appl. Phys.* 90, 5202 (2001).
399. W. H. Rippard, A. C. Perrella, and R. A. Buhrman, *Appl. Phys. Lett.* 78, 1601 (2001).
400. W. H. Rippard, A. C. Perrella, and R. A. Buhrman, *J. Appl. Phys.* 89, 6642 (2001).
401. B. Oliver, Q. He, X. Tang, and J. Nowak, *J. Appl. Phys.* 91, 4348 (2002).
402. H. Kikuchi, M. Sato, and K. Kobayashi, *Fujitsu Sci. Tech. J.* 37, 183 (2001).
403. M. Sato and K. Kobayashi, *IEEE Trans. Magn.* 33, 3553 (1997).
404. S. S. P. Parkin, K.-S. Moon, K. E. Pettit, D. J. Smith, R. E. Dunin-Borkowski, and M. R. McCartney, *Appl. Phys. Lett.* 75, 543 (1999).
405. R. C. Sousa, J. J. Sun, V. Soares, P. P. Freitas, A. Kling, M. F. da Silva, and J. C. Soares, *J. Appl. Phys.* 85, 5258 (1999).
406. J. Nowak, D. Song, and E. Murdock, *J. Appl. Phys.* 87, 5203 (2000).
407. Y. Ando, H. Kubota, M. Hayashi, M. Kamijo, K. Yaoita, A. Chak Chung Yu, X.-F. Han, and T. Miyazaki, *Jpn. J. Appl. Phys.* 39, 5832 (2000).
408. J. Schmalhorst, H. Brückl, G. Reiss, G. Gieres, and J. Wecker, *J. Appl. Phys.* 91, 6617 (2002).
409. T. Dimopoulos, C. Tiusan, V. da Costa, K. Ounadjela, and H. A. M. van den Berg, *Appl. Phys. Lett.* 77, 3624 (2000).
410. D. J. Keavney, S. Park, C. M. Falco, and J. M. Slaughter, *Appl. Phys. Lett.* 78, 234 (2001).
411. Z. Zhang, S. Cardoso, P. P. Freitas, P. Wei, N. Barradas, and J. C. Soares, *Appl. Phys. Lett.* 78, 2911 (2001).
412. Z. Zhang, S. Cardoso, P. P. Freitas, X. Batlle, P. Wei, N. Barradas, and J. C. Soares, *J. Appl. Phys.* 89, 6665 (2001).
413. T. Dimopoulos, V. Da Costa, C. Tiusan, K. Ounadjela, and H. A. M. van den Berg, *Appl. Phys. Lett.* 79, 3110 (2001).
414. N. Matsukawa, A. Odagawa, Y. Sugita, Y. Kawashima, Y. Morinaga, M. Satomi, M. Hiramoto, and J. Kuwata, *Appl. Phys. Lett.* 81, 4784 (2002).
415. M. G. Samant, J. Lüning, J. Stöhr, and S. S. P. Parkin, *Appl. Phys. Lett.* 76, 3097 (2000).
416. C. S. Yoon, J. H. Lee, H. D. Jeong, C. K. Kim, J. H. Yuh, and R. Haasch, *Appl. Phys. Lett.* 80, 3976 (2002).
417. Y. Saito, M. Amano, K. Nakajima, S. Takahashi, and M. Sagoi, *J. Magn. Magn. Mater.* 223, 293 (2001).
418. J. H. Lee, S. J. Kim, C. S. Yoon, C. K. Kim, B. G. Park, and T. D. Lee, *J. Appl. Phys.* 92, 6241 (2002).
419. M. A. M. Gijs, J. B. Giesbers, S. K. J. Lenczowski, and H. H. J. M. Janssen, *Appl. Phys. Lett.* 63, 111 (1993).
420. S. K. J. Lenczowski, R. J. M. van de Veerdonk, M. A. M. Gijs, J. B. Giesbers, and H. H. J. M. Janssen, *J. Appl. Phys.* 75, 5154 (1994).
421. R. J. M. vandeVeerdonk, J. Nowak, R. Meservey, J. S. Moodera, and W. J. M. deJonge, *Appl. Phys. Lett.* 71, 2839 (1997).
422. K. Matsuda, N. Watari, A. Kamijo, and H. Tsuge, *Appl. Phys. Lett.* 77, 3060 (2000).
423. J. S. Moodera, L. R. Kinder, J. Nowak, P. LeClair, and R. Meservy, *Appl. Phys. Lett.* 69, 708 (1996).
424. J. Chen, Y. Li, J. Nowak, and J. Fernandez de-Castro, *J. Appl. Phys.* 91, 8783 (2002).
425. H. Kikuchi, M. Sato, and K. Kobayashi, *Fujitsu Sci. Tech. J.* 37, 183 (2001).
426. K. B. Li, Y. H. Wu, J. J. Qiu, G. C. Han, Z. B. Guo, and T. C. Chong, *J. Magn. Magn. Mater.* 241, 89 (2002).
427. J. S. Moodera, L. R. Kinder, T. M. Wong, and R. Meservey, *Phys. Rev. Lett.* 74, 3273 (1995).
428. J. S. Moodera, J. Nowak, and R. J. M. van de Veerdonk, *Phys. Rev. Lett.* 80, 2941 (1998).
429. A. C. Marleyand and S. S. P. Parkin, *J. Appl. Phys.* 81, 5526 (1997).
430. J. Zhang and R. M. White, *J. Appl. Phys.* 83, 6512 (1998).
431. Y. Lu, X. W. Li, G. Xiao, R. A. Altman, W. J. Gallagher, A. Marley, K. Roche, and S. Parkin, *J. Appl. Phys.* 83, 6515 (1998).
432. R. S. Beech, J. Anderson, J. Daughton, B. A. Everitt, and D. Wang, *IEEE Trans. Magn.* 32, 4713 (1996).
433. N. Tezuka and T. Miyazaki, *Jpn. J. Appl. Phys.* 37, L218 (1998).
434. J. J. Sun and P. P. Freitas, *J. Appl. Phys.* 85, 5264 (1999).

435. C. Heide, A. I. Krikunov, Yu. F. Ogrin, P. E. Zilberman, and R. J. Elliott, *J. Appl. Phys.* 87, 5221 (2000).
436. G. G. Cabrera and N. García, *Appl. Phys. Lett.* 80, 1782 (2002).
437. X. H. Xiang, T. Zhu, J. Du, G. Landry, and J. Q. Xiao, *Phys. Rev. B* 66, 174407 (2002).
438. S. Zhang, P. M. Levy, A. C. Marley, and S. S. P. Parkin, *Phys. Rev. Lett.* 79, 3744 (1997).
439. Y. Saito, M. Amano, K. Nakajima, S. Takahashi, M. Sagoi, and K. Inomata, *Jpn. J. Appl. Phys.* 39, L1035 (2000).
440. Y. Saito, M. Amano, K. Nakajima, S. Takahashi, and M. Sagoi, *J. Magn. Magn. Mater.* 223, 293 (2001).
441. B. G. Park and T. D. Lee, *Appl. Phys. Lett.* 81, 2214 (2002).
442. C. H. Shang, J. Nowak, R. Jansen, and J. S. Moodera, *Phys. Rev. B* 58, R2917 (1998).
443. U. Rüdiger, R. Calarco, U. May, K. Samm, J. Hauch, H. Kittur, M. Sperlich, and G. Güntherodt, *J. Appl. Phys.* 89, 7573 (2001).
444. T. Hagler, R. Kinder, and G. Bayreuther, *J. Appl. Phys.* 89, 7570 (2001).
445. J. Wingbermühle, S. Stein, and H. Kohlstedt, *J. Appl. Phys.* 92, 7261 (2002).
446. J. H. Lee, K. I. Lee, W. L. Lee, K.-H. Shin, J. S. Lee, K. Rhie, and B. C. Lee, *J. Appl. Phys.* 91, 7956 (2002).
447. R. Coehoorn, S. R. Cumpson, J. J. M. Ruigrok, and P. Hidding, *IEEE Trans. Magn.* 35, 2586 (1999).
448. K. Shimazawa, O. Redon, N. Kasahara, J. J. Sun, K. Sato, T. Kagami, S. Saruki, T. Umehara, Y. Fujita, J. Yamazaki, S. Araki, H. Morita, and M. Matsuzaki, *IEEE Trans. Magn.* 36, 2542 (2000).
449. D. Song, J. Nowak, R. Larson, P. Kolbo, and R. Chellew, *IEEE Trans. Magn.* 36, 2545 (2000).
450. M. K. Ho, C. H. Tsang, R. E. Fontana, Jr., S. S. P. Parkin, K. J. Carey, P. Tao, S. MacDonald, P. C. Arnett, and J. O. Moore, *IEEE Trans. Magn.* 37, 1691 (2001).
451. S. Araki, K. Sato, T. Kagami, S. Saruki, T. Uesugi, N. Kasahara, T. Kuwashima, N. Ohta, J. Sun, K. Nagai, S. Li, N. Hachisuka, H. Hatate, T. Kagotani, N. Takahashi, K. Ueda, and M. Matsuzaki, *IEEE Trans. Magn.* 38, 73 (2002).
452. K. Machida, N. Hayashi, Y. Miyamoto, T. Tamaki, and H. Okuda, *J. Magn. Magn. Mater.* 201, 235 (2001).
453. K. Ishihara, M. Nakada, E. Fukami, K. Nagahara, H. Honjo, and K. Ohashi, *IEEE Trans. Magn.* 37, 1687 (2001).
454. K. Inomata and H. Sakakima, in "Giant Magneto-Resistance Devices" (E. Hirota, H. Sakakima, and K. Inomata), p. 133. Springer, Berlin, 2002.
455. K. Shimazawa, J. J. Sun, N. Kasahara, K. Sato, T. Kagami, S. Saruki, O. Redon, Y. Fujita, T. Umehara, J. Syoji, S. Araki, and M. Matsuzaki, *IEEE Trans. Magn.* 37, 1684 (2001).
456. J. C. Jury and S. X. Wang, *IEEE Trans. Magn.* 38, 295 (2002).
457. S. Ingvarsson, G. Xiao, S. S. P. Parkin, W. J. Gallagher, G. Grinstein, and R. H. Koch, *Phys. Rev. Lett.* 85, 3289 (2000).
458. E. R. Nowak, M. B. Weissman, and S. S. P. Parkin, *Appl. Phys. Lett.* 74, 600 (1999).
459. S. Ingvarsson, G. Xiao, R. A. Wanner, P. Trouilloud, Y. Lu, W. J. Gallagher, A. Marley, K. P. Roche, and S. S. P. Parkin, *J. Appl. Phys.* 85, 5270 (1999).
460. D. S. Reed, C. Nordman, and J. M. Daughton, *IEEE Trans. Magn.* 37, 2028 (2001).
461. K. S. Kim, H. J. Shim, I. J. Hwang, B. K. Cho, J. H. Seok, and J.-T. Kim, *J. Appl. Phys.* 91, 8804 (2002).
462. P. K. George, Y. Wu, R. M. White, E. Murdock, and M. Tondra, *Appl. Phys. Lett.* 80, 682 (2002).
463. N. Smith and P. Arnett, *Appl. Phys. Lett.* 78, 1448 (2001).
464. S. Mao, J. Nowak, E. Linville, H. Chen, O. Heinonen, H. Cho, L. Wang, P. Anderson, M. Ostrowski, B. Karr, P. Kolbo, and Z. Zhang, "47th Annual Conference on Magnetism and Magnetic Materials," Tampa, FL, Nov. 2002, paper DC-01.
465. Y. Irie, H. Sakakima, M. Satomi, and Y. Kawawake, *Jpn. J. Appl. Phys.* 34, L415 (1995).
466. J. M. Dauton, *J. Appl. Phys.* 81, 3758 (1997).
467. R. C. Sousa, P. P. Freitas, V. Chu, and J. P. Conde, *IEEE Trans. Magn.* 35, 2832 (1999).
468. J. M. Slaughter, R. W. Dave, M. DeHerrera, M. Durlam, B. N. Engel, J. Janesky, N. D. Rizzo, and S. Tehrani, *J. Supercond.* 15, 19 (2002).
469. S. Tehrani, B. Engel, J. M. Slaughter, E. Chen, M. DeHerrera, M. Durlam, P. Naji, R. Whig, J. Janesky, and J. Calder, *IEEE Trans. Magn.* 36, 2752 (2000).
470. D. Wang, M. Tondra, A. V. Pohm, C. Nordman, J. Anderson, J. M. Daughton, and W. C. Black, *J. Appl. Phys.* 87, 6385 (2000).
471. J. G. Zhu, Y. F. Zheng, and G. A. Prinz, *J. Appl. Phys.* 87, 6668 (2000).
472. K. Inomata, Y. Saito, K. Nakajima, and M. Sagoi, *J. Appl. Phys.* 87, 6064 (2000).
473. H. Boeve, C. Bruynseraede, J. Das, K. Dessein, G. Borghs, J. De Boeck, R. C. Sousa, L. V. Melo, and P. P. Freitas, *IEEE. Trans. Magn.* 35, 2820 (1999).
474. Y. Irie, H. Sakakima, M. Satomi, and Y. Kawawake, *Jpn. J. Appl. Phys.* 34, L415 (1995).
475. Y. Zheng and J.-G. Zhu, *IEEE Trans. Magn.* 33, 3286 (1997).
476. J. G. Zhu and Y. F. Zheng, *IEEE Trans. Magn.* 34, 1063 (1998).
477. B. A. Everitt, A. V. Pohm, R. S. Beech, A. Fink, and J. M. Daughton, *IEEE Trans. Magn.* 34, 1060 (1998).
478. B. A. Everitt, A. V. Pohm, and J. M. Daughton, *J. Appl. Phys.* 81, 4020 (1997).
479. R. R. Katti, *J. Appl. Phys.* 91, 7245 (2002).
480. H. Boeve, J. Das, C. Bruynseraede, J. De Boeck, and G. Borghs, *Electron. Lett.* 34, 1754 (1998).
481. T. N. Fang and J. G. Zhu, *IEEE Trans. Magn.* 37, 1963 (2001).
482. P. Shang, A. K. Petford-Long, and T. C. Anthony, *J. Appl. Phys.* 91, 7703 (2002).
483. J. Gadbois and J.-G. Zhu, *J. Appl. Phys.* 34, 1066 (1998).
484. J. G. Zhu and Y. F. Zheng, *Top. Appl. Phys.* 83, 289 (2002).
485. K. J. Kirk, J. N. Chapman, and C. D. W. Wilkinson, *Appl. Phys. Lett.* 71, 539 (1997).
486. K. J. Kirk, J. N. Chapman, and C. D. W. Wilkinson, *J. Appl. Phys.* 85, 5237 (1999).
487. G. Yi, P. R. Aitchison, W. D. Doyle, J. N. Chapman, and C. D. W. Wilkinson, *J. Appl. Phys.* 92, 6087 (2002).
488. F. J. Castano, Y. Hao, S. Haratani, C. A. Ross, B. Vogeli, M. Walsh, and H. I. Smith, *IEEE Trans. Magn.* 37, 2073 (2001).
489. J. Janesky, N. D. Rizzo, L. Savtchenko, B. Engel, J. M. Slaughter, and S. Tehrani, *IEEE Trans. Magn.* 37, 2052 (2001).
490. S. E. Russek, J. O. Oti, and Y. K. Kim, *J. Magn. Magn. Mater.* 199, 6 (1999).
491. M. Ooishi, *Nikkei Electron.* 839, 90 (2003) (in Japanese).
492. S. Tehrani, M. Durlam, J. Åkerman, M. DeHerrera, J. Salter, J. Slaughter, J. Janesky, B. Engel, and N. Rizzo, *Nikkei Electron.* 835, 127 (2002) (in Japanese).
493. R. C. Sousa, Z. Zhang, and P. P. Freitas, *J. Appl. Phys.* 91, 7700 (2002).
494. K. Inomata, Y. Saito, K. Nakajima, and M. Sagoi, *J. Appl. Phys.* 87, 6064 (2000).
495. F. Z. Wang, *Appl. Phys. Lett.* 77, 2036 (2000).
496. Y. K. Zheng, X. Y. Wang, D. You, and Y. H. Wu, *Appl. Phys. Lett.* 79, 2788 (2001).
497. R. S. Beech, J. A. Anderson, A. V. Pohm, and J. M. Daughton, *J. Appl. Phys.* 87, 6403 (2000).
498. W. C. Jeong, B. I. Lee, and S. K. Joo, *J. Appl. Phys.* 85, 4782 (1999).
499. W. C. Jeong, B. I. Lee, and S. K. Joo, *IEEE Trans. Magn.* 34, 1069 (1998).
500. Y. K. Zheng, Y. H. Wu, Z. B. Guo, G. C. Han, K. B. Li, J. J. Qiu, H. Xie, and P. Luo, *IEEE Trans. Magn.* 38, 2850 (2002).
501. J. Slonczewski, *J. Magn. Magn. Mater.* 159, L1 (1996).
502. Ya. B. Bazaliy, B. A. Jones, and S. C. Zhang, *J. Appl. Phys.* 89, 6793 (2001).

503. M. Tsoi, A. G. M. Jansen, J. Bass, W.-C. Chiang, M. Seck, V. Tsoi, and P. Wyder, *Phys. Rev. Lett.* 80, 4281 (1998).
504. J.-E. Wegrowe, D. Kelly, Y. Jaccard, Ph. Guittienne, and J.-Ph. Ansermet, *Europhys. Lett.* 45, 626 (1999).
505. J. Z. Sun, *J. Magn. Magn. Mater.* 202, 157 (1999).
506. E. B. Myers, D. C. Ralph, J. A. Katine, R. N. Louie, and R. A. Buhrman, *Science* 285, 867 (1999).
507. E. B. Myers, D. C. Ralph, J. A. Katine, F. J. Albert, and R. A. Buhrman, *J. Appl. Phys.* 87, 5502 (2000).
508. J. Grollier, V. Cros, A. Hamzic, J. M. George, H. Jaffrès, A. Fert, G. Faini, J. Ben Youssef, and H. Legall, *Appl. Phys. Lett.* 78, 3663 (2001).
509. W. W. Zhuang, W. Pan, B. D. Ulrich, J. J. Lee, L. Stecker, A. Burmaster, D. R. Evans, S. T. Hsu, M. Tajiri, A. Shimaoka, K. Inoue, T. Naka, N. Awaya, K. Sakiyama, Y. Wang, S. Q. Liu, N. J. Wu, and A. Ignatiev, "2002 International Electron Devices Meeting Technical Digest," San Francisco, CA, Sept. 2002, pp. 193–196.
510. H. Ohno, D. Chiba, F. Matsukura, T. Omiya, E. Abe, T. Dietl, Y. Ohno, and K. Ohtani, *Nature* 408, 944 (2000).
511. S. Datta and B. Das, *Appl. Phys. Lett.* 56, 665 (1990).
512. D. J. Monsma, J. C. Lodder, Th. J. A. Popma, and B. Dieny, *Phys. Rev. Lett.* 74, 5260 (1995).
513. T. Hayashi, H. Shimada, H. Shimizu, and M. Tanaka, *J. Cryst. Growth* 201/202, 689 (1999).
514. D. Chiba, N. Akiba, F. Matsukura, Y. Ohno, and H. Ohno, *Appl. Phys. Lett.* 77, 1873 (2000).
515. J. Fabian, I. Žutić, and S. D. Sarma, *Phys. Rev. B* 66, 165301 (2002).
516. M. Johnson, B. R. Bennett, M. J. Yang, M. M. Miller, and B. V. Shanabrook, *Appl. Phys. Lett.* 71, 974 (1997).
517. S. van Dijken, X. Jiang, and S. S. P. Parkin, *Appl. Phys. Lett.* 80, 3364 (2002).
518. R. Fiederling, M. Keim, G. Reuscher, W. Ossau, G. Schmidt, A. Waag, and L. W. Molenkamp, *Nature* 402, 787 (1999).
519. Y. Ohno, D. K. Young, B. Beschoten, F. Matsukura, H. Ohno, and D. D. Awschalom, *Nature* 402, 790 (1999).
520. B. T. Jonker, Y. D. Park, B. R. Bennett, H. D. Cheong, G. Kioseoglou, and A. Petrou, *Phys. Rev. B* 62, 8180 (2000).
521. N. García, M. Muñoz, G. G. Qian, H. Rohrer, I. G. Saveliev, and Y.-W. Zhao, *Appl. Phys. Lett.* 79, 4550 (2001).
522. S. Z. Hua and H. Deep Chopra, *Phys. Rev. B* 67, 060401 (2003).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanostorage

Jooho Kim

Samsung Electronics, Suwon, Korea

CONTENTS

1. INTRODUCTION

For terabyte memory that will be necessary after five–ten years in the information industry, nanostorage may be the most promising solution, together with three-dimensional memory using a hologram method. In general, nanotechnology means a technology able to make a structure below the 100 nm scale with good efficiency. It is defined that nanostorage also means information storage of nanometer scale, which has below 100 nm data bits. In order to retrieve the below 100 nm data bits, we should use near-field optics or surface plasmon phenomena with a current optical system. First, as a typical method, a superresolution near-field structure will be described. Next, we will give an account of an electrical nanostorage method. Other technologies such as scanning near-field optical microscopy (SNOM), solid immersion lens (SIL), magnetic amplifying magneto-optical system (MAMMOS), and heat-assisted magnetic recording (HAMR) for nanostorage also will be explained briefly. Although each technology has a potential, it is thought that a terabyte memory may be achieved much faster through the fusion of these technologies.

2. NANOSTORAGE BY SUPERRESOLUTION NEAR FIELD

A near field represents the electromagnetic field near the surface of a fine probe, a hole, or scattering center such as silver (Ag), gold (Au), aluminum (Al), and so on. Near the surface, there exist very short wavelength components, compared with that of incident light. These short-wavelength components are able to retrieve the signals below the resolution limit. Through conventional far-field optics, we can write and retrieve the signal just within the resolution limit, $\lambda/(4 \cdot \mathrm{NA})$, where λ is the laser wavelength and NA is the numerical aperture of the optical system. From current study of a system of 405 nm (λ) and 0.85 (NA), we can write and retrieve around 150 nm mark length signals by a phase-change media storage system, that is the so-called Blu-Ray disk [1]. This resolution limit theory is also applied to a magneto-optical storage system. Basically, we cannot retrieve signals of below 100 nm from a current optical system. Therefore, methods using an SIL and an SNOM have been studied for retrieving magneto-optical (MO) signals by near-field coupling thus far. By the SNOM near-field technology, it has been reported that signals below 100 nm can be retrieved [2–8]. Recently, the superresolution technique combined with the near field also has been introduced [9, 10]. This system has been studied with phase change recording materials such as GeSbTe, AgInSbTe, and surface plasmon near-field signals have been analyzed by the reflectivity difference caused by the phase change. Within a superresolution near-field (SRNF) structure, a specially designed mask layer is partially changed to a transparent (Sb) or light-scattering center ($\mathrm{Ag} + \mathrm{O}_x$) by laser heat distribution. This partially changed area of the mask layer generates an aperture-type (Sb) near field or a scattering center (surface plasmon)-type (AgO_x) near field. In this section, we introduce a variety of recording systems in the superresolution technique combined with the near field. First, we will discuss whether magneto-optical signals also keep their Kerr

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (545–552)

polarization feature in the surface plasmon near field, and can be applied to the superresolution near-field structure. As another topic, the reactive diffusion recording mechanism and its superresolution technique using a near field is described in order to improve the resolution limit. The superresolution near-field structure medium has the almost same principle and results in both a magneto-optical and a phase change medium. Results using typical magneto-optical media will be described in the next section.

2.1. Magneto-Optical/Phase Change Nanostorage

In MO disks, magnetic superresolution (MSR) and MAMMOS techniques have been proposed with special readout and recording magnetic layers consisting of GdFeCo and TbFeCo [11, 12]. And it has been reported recently that a transparent aperture (Sb)-type superresolution near-field structure MO disk can also enhance the signal intensity of the below resolution limit [13]. Moreover the light-scatter-center-type (AgO_x) superresolution near-field MO disk overcomes the theoretical resolution limit, and can retrieve signals up to a 100 nm mark length. A carrier-to-noise ratio (CNR) at a mark length of 150 nm, which is below the resolution limit, is more than 20 dB compared to 0 dB in conventional disks. This means that the Ag particles or their clusters generated in AgO_x film can enhance MO signals by approximately 100 times those of the conventional MO disks, and a nonmagnetic light-scattering center can conserve the polarized MO signals. This optical mask layer for near-field retrieval is nonmagnetic, but has optically nonlinear characteristics. The basic structure of a typical AgO_x superresolution disk is composed of six layers: dielectric ($ZnS–SiO_2$, 60 nm)/mask (AgO_x, 15 nm)/dielectric ($ZnS–SiO_2$, 25 nm)/recording (TbFeCo, 25 nm)/dielectric ($ZnS–SiO_2$, 25 nm)/reflective (Ag, 50 nm) layers. All films are deposited by RF magnetron sputtering with the composite targets, except for the AgO_x film. The AgO_x film is produced by RF-reactive magnetron sputtering using a pure Ag target and a gas mixture of oxygen (O_2) and argon (Ar).

To obtain the superresolution near-field effect in MO disks, we determine the characteristics of the superresolution near-field MO disk with an AgO_x mask film at different readout powers [13]. Signal enhancement is obtained up to a 100 nm mark length at a high readout power of 3.0 mW. In contrast, the same disks do not show any signal enhancement at a low readout power of 1.0 mW. This suggests that the AgO_x layer changes partially to heat-generated Ag by strong laser heat, as reported by Tominaga et al. [14]. Although a MAMMOS technique can amplify crescent-shaped domains of 80 nm breadth, MO signals from circular magnetic domains with a size of 300 nm cannot be enhanced by applying MAMMOS [15]. However, it has been reported that the circular domains can be detected by near-field coupling up to less than 100 nm using an optically nonlinear AgO_x mask layer. The circular shape can be confirmed using a magnetic-force microscope (MFM) [13]. In MAMMOS, the length of the crescent-shaped magnetic domains in the radial direction is usually about five times more than that of the breadth. It is thus difficult to increase the radial recording density. On the other hand, the marks recorded with the superresolution near-field technique are circular, and the area density is much larger than that in MAMMOS. To examine the wavelength effect, authors measured the signal intensities of one sample disk using two different disk drives with a wavelength of 680 nm (NA 0.55) or 405 nm (NA 0.65). In the case of a 680 nm wavelength, they could read out up to a 170 nm mark length. On the other hand, in the case of a 405 nm wavelength, the resolution was up to a below 100 nm mark length. On the basis of this result, a shorter wavelength also retrieves much smaller marks in near-field optics. It has also been reported that the conservation ability of a polarized MO signal in an AgO_x light scattering type is much higher than that of Sb transparent aperture-type superresolution near-field MO disks from carrier to noise ratio (CNR) comparison.

To reveal signal enhancement by the light-scattering center in superresolution near-field disks, the electric-field intensity has been estimated using Sb and AgO_x as the mask layer by computer simulation with a finite-differential time-domain (FDTD) method, as shown in Figure 1 [13]. The light-scattering center and transparent aperture were assumed as an Ag particle and as an Sb amorphous transparent aperture. The aperture and scattering center were fixed at 200 nm in this simulation. It has been reported that the z component of the electric field of the light-scattering center (AgO_x) is about 20 times stronger than that of the transparent aperture (Sb) underneath the mask layers. This result agrees with the experimental results, although polarization components of the MO layer are not considered in this simulation. As is well known, a surface plasmon only appears under some conditions at the interface between a metal and a dielectric material. In order to generate the surface plasmon effectively, we have to use a total reflection method, or gratings, or a rough surface in general [16]. Shalaev describes that the aggregated Ag particles or their clusters with a size range of $\sim$100 nm may also generate plasmons, resulting in huge local-field enhancements of optical nonlinear factors in some special sizes and its fractal dimensions [17]. The strong electric-field enhancement, in some cases, can be up to 10^{10}, which is four orders magnitude larger than the average for the unique film. The enhancement effect was actually observed as surface-enhanced Raman images by NSOM. On the other hand, by the refractive-index evaluation of AgO_x in recent work, the index is shifted from its initial AgO_x (mostly Ag_2O) with $2.5+1.5j$ to a more Ag-rich state with $0.7+3.7j$ as the readout power is increased, generating a light-scattering center. The Ag-rich local phase may be compatible with the localized states of Ag mentioned by Shalaev [17]. It should be noted that the electric field returned or transmitted from superresolution near-field structure disks is partly polarized. This effect plays an important role for signal preservation in superresolution near-field MO signals, although details are not clear at the moment. It has been confirmed that the 100 nm MO signal can be retrieved by the superresolution near-field effect with a light-scattering center (AgO_x), and it has been found that the nonmagnetic layer can conserve the Kerr polarization. The results indicate the possibility to enhance the resolution of MO circular signals further, up to less than a 100 nm mark length.

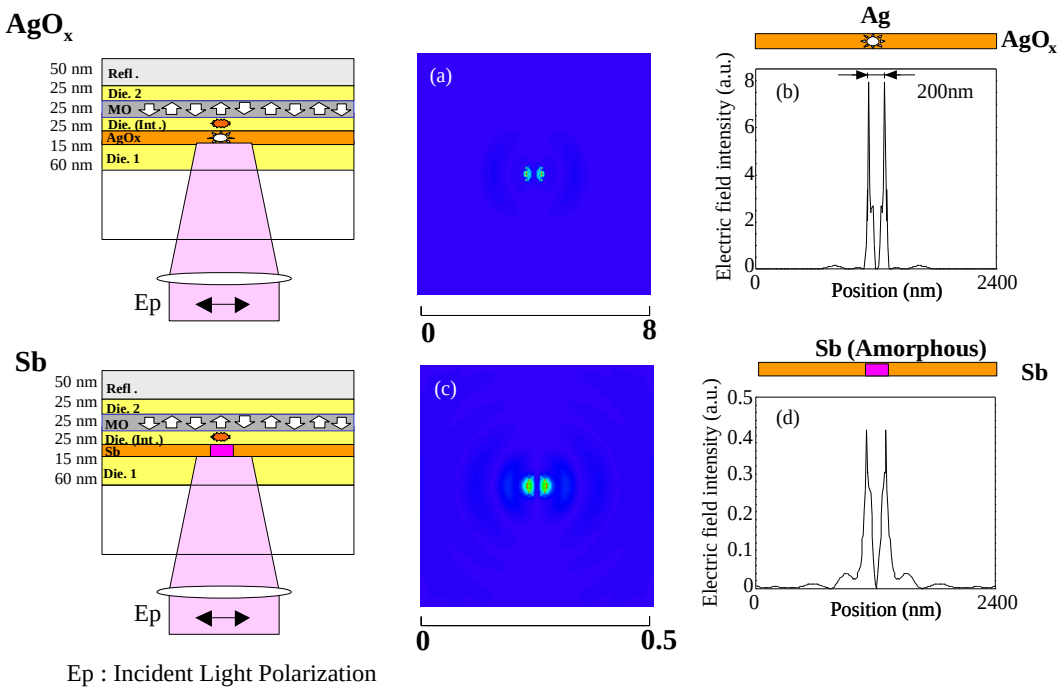


Figure 1. Electric fields of z components beneath mask layers calculated by FDTD. (a) Near-field distribution scattered in the X–Y plane of superresolution near-field structure using AgO_x film and (b) field intensity. (c) Near-field distribution scattered in the X–Y plane of superresolution near-field structure using Sb film and (d) field intensity. Reprinted with permission from [13], J. H. Kim et al., *Appl. Phys. Lett.* 77, 1774 (2000). © 2000, American Institute of Physics.

2.2. Thermal Reactive Diffusion Nanostorage

The recording mechanism of MO data storage is magnetization reversal by an external magnetic field and laser heat (thermomagnetic recording), which results in magnetic Kerr and Faraday rotation of incident light [18]. In contrast, phase change (PC) recording uses the transformation from the amorphous to crystalline phase by external laser heat, which results in a reflectivity change [19]. To achieve terabyte capacity data storage, a superresolution near-field structure has been proposed, and succeeded in retrieving signals less than 100 nm mark length using an AgO_x (light-scattering center type) or Sb (transparent aperture type) mask layer [20, 21]. As a recent result, it has been proposed to carry out reactive recording by applying a rare-earth transition metal (RE–TM) for making the nanostorage, and the reaction process and mechanism have been reported. [22]. To elucidate the mechanism, authors examined the optical constants, magnetic, and thermo-optical properties. In this section, a reactive recording with RE–TM and application to optical data storage will be described for 50 nm nanostorage.

The structure has been designed by considering the crystallization of RE–TM and the reaction between RE–TM and dielectric layers. To reveal the reaction process and mechanism, the thermo-optical property of the RE–TM samples having a three-layer structure sandwiched between ZnS–SiO_2 dielectric layers has been examined. The transmittance of the as-deposited and 783 K heated samples greatly changed from 400 to 1300 arbitrary units; the difference was about 900. This behavior is similar to a completely oxidized RE–TM single-layer sample. In the case of the phase change recording material (PC, GeSbTe) sample, the difference was just 150. Comparing the as-deposited sample to the 873 K heated sample, optical constants of RE–TM film at a 633 nm wavelength, n, k (refractive index: n, extinction coefficient: k) were changed from 3.3, 4.0 (as-deposited) to 2.2, 0.3 (873 K heated), respectively. This means that the sample becomes almost transparent. It is thought that the above change is related to the crystallization and reaction of RE–TM. It has been reported that RE–TM is first crystallized to α-TM and α-RE, and then followed by crystallization of the intermetallic compound of RE–TM in amorphous RE–TM film [23]. It is supposed that the reaction includes both the sulfuration and oxidation of RE–TM (mainly sulfuration). It has been reported that sulfur in ZnS–SiO_2 is easily released, and diffuses into a PC (GeSbTe) recording layer by external heat, then significantly changes the optical constants [24, 25]. Oxidation of amorphous RE–TM has also been mentioned by several researchers; the preferential oxidation is of a Tb atom (strong affinity for oxygen) and an oxidation reaction with SiO_2 dielectric layers of RE–TM [26–28]. It is thought that the crystallization of amorphous RE–TM film occurs predominantly during the range of the first step (573–763 K) as a result of a slow increase of transmittance. In the region of the second step (763–783 K), the thermal-activated reaction of the film seems to occur predominantly as a result of the rapid increase of transmittance. An attempt was also made to obtain the activation energy for the transition from Kissinger's formula ($\ln(A/T_t^2) = E_a/k_B T_t + C$), where A is the heating rate, T_t is the transition temperature at heating rate A, and k_B is the Boltzmann constant [29]. The activation energy for the transition of the sample was

about 2.31 eV, which is similar to the 2.51 eV of the GeSbTe phase change material [30]. There is no large increase of transmittance in an RE–TM film sandwiched with Si_3N_4 dielectric layers with increasing temperature. This means that free sulfur and nonbridging oxygen (NBO) within RF-sputtered ZnS–SiO_2 dielectric layers plays an important role in the reactive recording of an RE–TM film [31]. Examining the photograph taken by a transmission electron microscope (TEM), it has been found that the RE–TM thin film is crystallized and reacted; the RE–TM layer is diffused into ZnS–SiO_2 dielectric layers, and volume expansion occurred in the 783 K heated sample. Authors have also investigated the magnetic properties of the samples to confirm the reacted state of the RE–TM film. The magnetic properties (coercivity: *Hc*, magnetization: *Bs*) of the heated-sample (873 K) sandwiched with ZnS–SiO_2 dielectric layers were the same as those of completely oxidized RE–TM film. The heated sample (873 K) sandwiched with Si_3N_4 dielectric layers, however, kept the magnetic properties (lower *Hc* and larger *Bs* than those of the as-deposited sample). From this result, it has also been confirmed that the oxidation reaction of the RE–TM recording layer occurs with increasing temperature. All of these reactions make it possible to store nanoscale data bits of around 50 nm.

Figure 2 represents the basic apparatus for optical storage and characteristics of the sample. The RE–TM recording layers have been applied to a superresolution near-field structure disk, and compared with a PC superresolution near-field structure disk with an AgO_x mask layer. The CNR (24 dB) of the sample disk by RE–TM (MO) reactive recording was twice as high as the result (12 dB) of the PC superresolution near-field structure disk for 100 nm mark length signals. Examining the spectrophotometer results, it was found that the samples of three-layer structure have a similar large reflectivity change (20%) in the blue laser region (405 nm), compared with 23% in the red laser region (635 nm) between as-deposited and 873 K heated samples. It is thus thought that we can apply the sample disks to the blue laser drive system as well as the red drive system, and we can expect to have better CNRs in the blue laser drive system. Current results using a blue laser (405 nm) system reveal the 50 nm signal retrieval, which has 15 dB CNR.

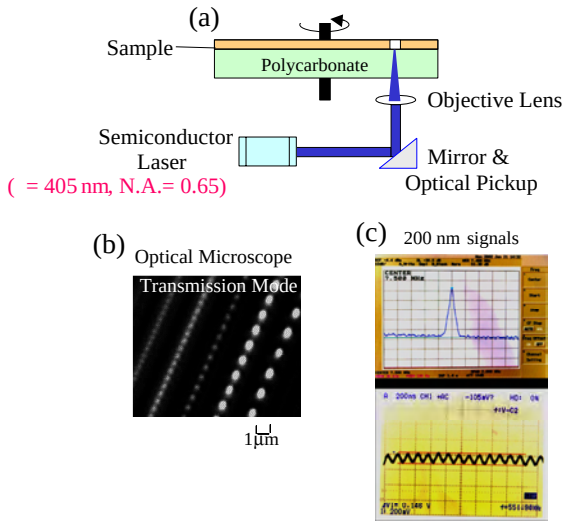


Figure 2. Sample preparation for optical application. (a) Laser application apparatus. (b) Optical microscope photograph fabricated by controlling the laser power. (c) Optical characteristics.

3. NANOSTORAGE BY REACTIVE ELECTRICAL CIRCUIT

The concept that nanometer level devices perform the basic functions of digital electronics may be realized by means of a variety of techniques, such as experiments based on monomolecular devices [32], molecular junctions [33, 34], electrodeposition of nanoparticles [35], and nanolithography [36]. In this section, we will introduce the attainment of complete electrical nanodevices based on the Hall effect through a thermal-activated reaction between rare-earth transition metal and dielectric material, using a nanosecond laser pulse or nanoresolution electron beam reactive electrical circuit (REC) method [37]. A physically mixed phase structure by cosputtering with RE–TM and dielectric material, which shows a metallic conductivity (above $10^3\ \Omega^{-1} \cdot cm^{-1}$) in less than 10% ZnS–$SiO_2$ composition, was used in order to generate the whole homogeneous reaction. The thermal-activated reaction and the interdiffusion related to a sulfuration and an oxidation of RE–TM [22] result in an outstanding change of electronic transport properties. The resistivity of the mixed structure samples is steeply altered by a magnitude of about 10^3 according to the increase of the reaction temperature, which may be strongly related to the previous thermal-activated reaction (Fig. 3). This result suggests that an electrical insulator and a conductor can be fabricated by controlling the reaction temperature. The 773 K heated sample decreases the resistivity with

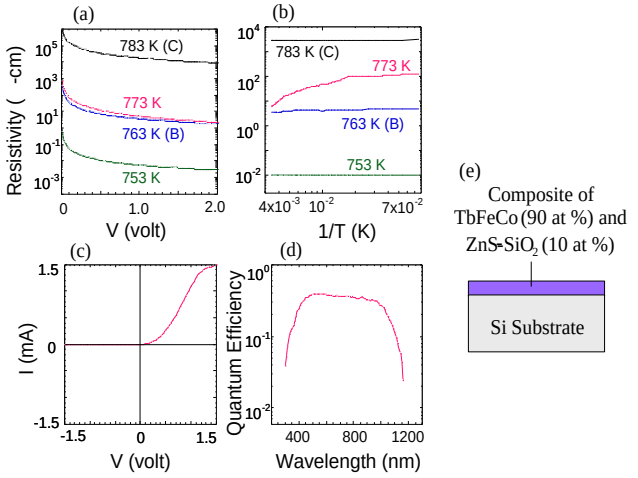


Figure 3. Electronic transport properties. Substrates in (a), (b) plots are quartz, and those in (c), (d) are Si wafer. (a) Effect of heating temperature on resistivity. (b) Temperature dependence of resistivity $R(T)$: the voltage was measured at a fixed current of around 0.1 μA. (c) I–V curve of the 773 K sample on an Si substrate (p-type). (d) Wavelength dependence of quantum efficiency: the cell size is 7×7 mm^2, and incident photons are 4.53×10^{13}/cm^2. (e) Structure and materials of the reactive semiconductor sample. Reprinted with permission from [37], J. H. Kim et al., *Appl. Phys. Lett.* 80, 2764 (2002). © 2002, American Institute of Physics.

increasing temperature like a typical semiconductor, while the other samples do not show such a behavior. A notable distinction among the samples is the degree of the oxidation, sulfuration, and diffusion of Fe, Co. In order to confirm the semiconductor properties more precisely, samples were prepared on an Si substrate, and the *p–n* junction characteristics were measured. The *I–V* (current–voltage) curve of the reacted sample at 773 K on Si substrate (*p*-type) shows a clear characteristic of the *p–n* junction diode, where V_F (forward voltage) is around 0.3 V. The quantum efficiency of the sample on Si substrate (*p–n* junction) as a photodiode is about 0.4 in the range of 450–1000 nm wavelength. This means that semiconductor devices such as a photodiode and a solar cell [38, 39] can be fabricated by the above simple thermal-activated reactions (Fig. 3).

The transparent thermal-activated reaction areas with a circular and line shape were confirmed by an optical microscope (transmittance mode) photograph. In the case of the circular shape, about a 50 nm size reacted dot was obtained, using an optical system with a laser wavelength (λ) of 405 nm and a lens numerical aperture of 0.65. The semiconductor or insulator sections are obtained by controlling the laser beam power intensity. The result represents the possibility of 50 nm dot or line fabrication. Moreover, if an electron beam is applied to the fabrication, it is expected that the 5–10 nm dots or lines [36] will be obtained by the above reaction.

A nanoelectronic device can be fabricated using the above reactions and the Hall effect. The Hall field $(\mathbf{E_H})_y$ [40] is the electric field $(\mathbf{j} \times \mathbf{B})_y$, developed across two faces of a (semi-) conductor when a current $\mathbf{j_x}$ flows across a magnetic field $\mathbf{B_z}$. By combining the incident current and external perpendicular magnetic field in the sample memory cell, an electronic signal will be put in storage within the semiconductor or condenser (insulator) area [37]. Authors, thus, have suggested that a 2.5×10^9–2.5×10^{11} (2.5–250 G) bit/in^2 capacity of flash memory or DRAM can be fabricated by means of the nanosecond laser pulse or electron-beam-induced reaction by some beam intensity profile control. A variety of devices, like a DRAM, flash memory, Hall element, *p–n* diode, and metal–oxide–semiconductor (MOS) transistor, may be fabricated with the combination of the insulator, semiconductor, and metal areas by controlling the electron beam or laser power intensity whose temperature can be calculated by computer simulation and actual experiments [41]. The results in this technology indicate the potential of a thermal-activated reaction to fabricate a variety of electronic devices for nanostorage.

4. NANOSTORAGE BY SCANNING NEAR-FIELD OPTICAL MICROSCOPE

Betzig and coworkers have tried to record and retrieve nanometer scale magneto-optical signals by an SNOM [42–45]. They obtained a 60 nm magneto-optical domain, which is equivalent to a 100 Gbit/in^2 areal density. An argon ion laser was used for reading and writing. A neutral density filter was placed in the read beam so that read powers were below the threshold for modifying the written bits. An acousto-optic modulator (AOM) switched the write beam on and off for selectively writing patterns into the film. Retardation plates were used to control the polarization state of the light at the aperture since the MO reading is strongly polarization dependent. The Faraday effect causes the input light to be rotated as it traverses the samples, with the direction of rotation dependent on the magnetization of the sample in any particular region (Fig. 4).

The sample was ten multilayers of alternating 4 Å Co and 10 Å Pt sputter coated onto a glass substrate. This material was sufficiently ferromagnetic to result in a permanent magnetization that will allow for observation of the Faraday effect. In this case, application of an external field is not necessary, and the polarity of the permanent magnetization can be modified locally to write bits into the sample. The SNOM was also used to write MO bits using the probe. By increasing the input power, the temperature of the sample immediately beneath of the tip was raised to the Curie temperature, which might be roughly 300 °C. This allows the magnetization of this heated region to switch its polarity to the opposite of its surrounding areas. For an input power up to 4.5 mW, no domain is formed. The authors suggest that either the power is too low for the sample to be heated near the Curie temperature or the heated region is too small to form a stable magnetic domain. They have succeeded in a domain formation of 60 nm size at higher powers such as 6.0–8.0 mW.

Authors have suggested that the bit density may increase to up to 100 Gbit/in^2 by using cleaner films (reducing bit errors) and using applied magnetic fields to increase the contrast, which is equivalent to 60 nm magnetic domain mark storage. Although the contrast mechanism responsible for the near-field writing process is not completely clear, the authors were convinced that it is an MO effect [42], that is, that it involves a rise to the Curie temperature and a flipping of the magnetic domain: whether through a thermal or an optical effect. Although write speeds in these experiments are limited to ~10 kHz, the authors suggest that ~10 MHz/channel may be possible. Obvious improvements to this technique might involve better film surface cleanliness and decreased fly height (to below 10 nm), in keeping with current directions in magnetic storage research. As a technique used with SNOM, but with potentially significant benefit to SNOM data storage, Betzig et al. also introduced the fiber laser probe [46]. With this probe, output power equivalent to about 10^{14} photons/s was achieved through tuning of the boundary conditions at a tapered near-field aperture that served as the back reflector of the fiber laser cavity. This represents an improvement of

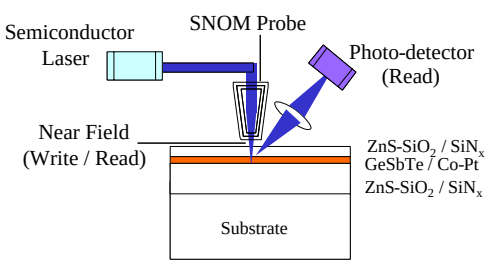


Figure 4. Schematic diagram of SNOM system and structure.

about three–four orders of magnitude over the photon output of conventional probes. Of particular importance in the present application are the higher read speeds that should result.

5. NANOSTORAGE BY SOLID IMMERSION LENS

The SIL has been introduced as a near-field instrument of optical microscopes, as shown in Figure 5. Terris and coworkers have utilized an SIL configuration for implementing near-field high-density MO data storage [47, 48]. The resolution limit of classical optics is inversely proportional to the numerical aperture (NA) of the optical system. Thus, increasing the NA decreases the distance between distinguishable features. Although the SIL approach appears to have a fundamental density limit, it has a benefit in significantly faster read and write rates that are much closer to those of the current MO technique. And although this SIL technique presents a near-field alternative to the SNOM methods for improving the density of MO data storage, significant obstacles remain. Perhaps most troubling is the necessity to provide a practical method of rapid flying lens above the MO medium at a height that keeps the lens in the near field of the medium, that is, ~50 nm. The authors suggested that if this could be accomplished, use of higher index SILs and shorter wavelength sources might result in 100 nm signal data retrieving [47, 48].

Kishima et al. [49] have developed a near-field optical storage using a superhemispherical SIL [47] in an attempt to improve the recording density. Therefore, in combination with a blue–violet light source of 405 nm wavelength, an areal density of over 50 Gbit/in^2 (80 nm bit signal) is expected. In the realization of near-field data storage, precise control of the air gap between the optical head and the recording medium is important, as described above. Authors have achieved active control of the air gap by developing a capacitance air-gap servo method detection of electrostatic capacitance by forming an electrode on the bottom of the SIL, and controlling the optical head position by mounting it on a two-axis actuator [50]. Furthermore, adopting a 1 mm diameter miniature SIL resulted in a lightweight lens unit and an improved servo bandwidth. The developed optical head had satisfactory optical characteristics, and enabled reading and writing with 80 nm bit signals. A phase change material, that does not required an external magnetic field, was used as a recording medium. Generally, the phase change medium consists of GeSbTe recording material sandwiched between dielectric ZnS–SiO_2 materials. In their work, near-field recording using the phase change medium with a land and groove structure was verified, and 80 nm mark storage was also confirmed with a TEM observation, a result which is equivalent to 50 Gbit/in^2 data storage.

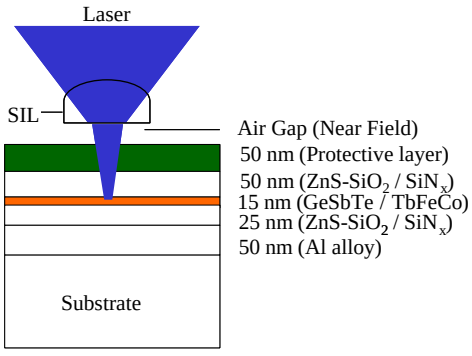


Figure 5. Structure and materials for SIL nanostorage.

6. NANOSTORAGE BY MAGNETIC AMPLIFYING MAGNETO-OPTICAL SYSTEM

Magneto-optical storage possesses a very great advantage of high-density recording compared to nonmagnetic optical storage because MO recording is almost the same as a perpendicular magnetic recording. However, the high-density recording below a 100 nm data bit is recorded by very tiny domains, so the readout MO signal decreases, and correct data reproduction becomes impossible. In order to amplify the small MO signal, MAMMOS technology has been proposed [51, 52]. The readout layer (GdFeCo) and recording layer (TbFeCo) were prepared by a magnetron sputtering process on a polycarbonate substrate of 1.2 mm thickness. Both magnetic layers were separated with nonmagnetic intermediate layers (SiN_x). The Curie temperature (T_c) and compensation temperature (T_{comp}) of the GdFeCo are about 350 °C and room temperature, respectively. Also, the corresponding values for the TbFeCo layers are about 270 °C and room temperature. The external field (H_e) in the expanding direction makes the size of the copied domain in the readout layer (GdFeCo) increase up to the same size as the laser spot. In this process, the readout signal is amplified to the signal saturation level, and is retrieved as shown in Figure 6 [51].

Authors have tried to detect much smaller MO domains than a 100 nm size. Pulse-width signals of a recording field from 20 to 100 ns were applied to a MAMMOS medium. These values can be estimated as a recording domain length of 20–100 nm since the disk linear velocity was 0.8 m/s. The signal waveform of the MAMMOS disk with a lower readout power and 200 Oe readout field shows a clear MO signal in the 20 nm domain, and the signal intensity is almost saturated. From this result, it was confirmed that a 20 nm length magnetic domain can be reproduced clearly, although the data transfer rate is a little late [53].

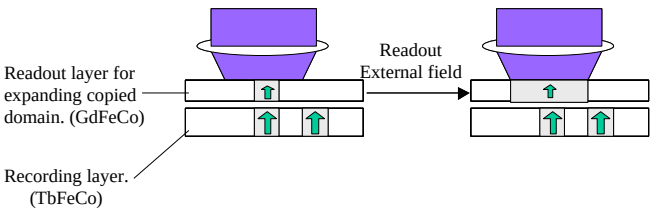


Figure 6. Domain expansion of MAMMOS disk.

Table 1. Basic characteristics of a variety of nanostorage methods.

	SRNF	REC	SNOM	SIL	MAMMOS	HAMR
Minimum bit length	50 nm	50 nm	60 nm	80 nm	40 nm × 80 nm	50 nm
Areal density	250 Gbits/in^2	250 Gbits/in^2	100 Gbits/in^2	50 Gbits/in^2	200 Gbits/in^2	250 Gbits/in^2
Hardware (application)	Next DVD	DRAM flash memory	Next DVD, MO	Next DVD, MO	Next MO	HDD
Principle	Near field	Hall effect	Near field	Near field	Amplifying domains	Heat-assisted magnetic recording
Materials	MO/PC	RE–TM	MO/PC	MO/PC	MO	Magnetic
Issue point	Signal intensity	Mass production	Head–media spacing	Head–media spacing	Data-transfer rate	Head/magnetic material

7. NANOSTORAGE BY HEAT-ASSISTED MAGNETIC RECORDING

To assure the thermal stability of the magnetic data, the medium must satisfy $K_uV/k_BT > 70$, where K_u is a uniaxial anisotropy constant, V is the isolated grain volume, k_B is Boltzmann's constant, and T is the ambient temperature [54]. This is commonly termed the superparamagnetic limit. A higher density (smaller bit cells) requires a larger media anisotropy, which in turn requires a large coercivity. The coercivity of the high-density medium approaches the maximum field that can be generated by conventional recording heads. It has been suggested that the superparamagnetic limit can be overcome by heating the media optically to momentarily reduce the coercivity of the media during recording. Very small aperture lens (VSAL) and SIL technologies have been suggested for generating extremely small spots [55–57]. The aperture in the shape of a rigid waveguide is deposited onto the end facet of a semiconductor laser. The waveguide is operated above cutoff, but FDTD calculations indicate that the rigid constriction amplifies the field in the center of the waveguide. A bow-tie antenna, in some sense the complement of the rigid waveguide, has also been proposed as a means for concentrating the field at the center of the antenna gap [58, 59]. This HAMR technology combined with above technologies has already demonstrated a magnetic spot size below 100 nm in hard disk storage technology [60]. Recently, using a giant magnetoresistive (GMR) head, a tunneling magnetoresistive (TMR) head, and perpendicular recording, it was reported that 250 Gbits/in^2–1 Tbit/in^2 would be able to be achieved [55, 61].

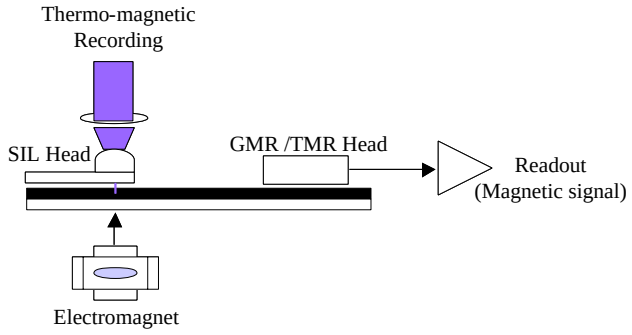


Figure 7. Recording and readout system of HAMR.

8. CONCLUSIONS

The basic characteristics of a variety of nanostorage methods are summarized in Table 1. Each technology has its originality and potential for terabyte memory or nanostorage, and is expected to be developed in three–five years. Although these technologies also have problems of commercialization due to the difficulties of a nanometer-scale mechanical control or a signal process of near-field light, a terabyte nanostorage actually should be introduced into the storage industry in five–seven years. At that time, all storage technologies may be fused into one, including the hard disk medium, which may be called a fusion memory or a synthesis memory using plural technologies.

GLOSSARY

Faraday and Kerr effects When polarized light transits through a magnetized material (Faraday effect) or reflects on the surface (Kerr effect), the polarized direction is altered by the effect of the magnetization.

Giant magnetoresistive (GMR) head Head perceiving the variation of electric resistance according to the magnetization direction of media. This head is designed to utilize a giant magnetoresistive phenomenon.

Hall effect An electric field is generated when a current flows across a magnetic field on a (semi-) conductor's two faces. The generated electric field is called the Hall field, and this phenomenon is called the Hall effect.

Optical mask layer Thin-film layer that became different optical constants by a laser intensity or a thermal reaction.

Superparamagnetic limit Magnetic-material size limit that changes from a ferromagnetic to a superparamagnetic material. Below the limit, ferromagnetic material loses its remanent magnetization in the state of no applied field (impossible to store magnetic data).

Surface plasmon near field (Localized) surface plasmon occurs by the vibration of free electrons on the tip of a metal probe as being irradiated by a laser or light. In this case, if the tip diameter is several tens of a nanometer, a spot size of several tens of a nanometer can be formed with much stronger light intensity on the near the tip, although incident light has a wavelength of several hundreds of a nanometer.

Tunneling magnetoresistive (TMR) head Head perceiving the variation of electric resistance according to the

magnetization direction of media. This head is designed to utilize a tunneling magnetoresistive phenomenon.

REFERENCES

1. M. Shoji, A. Nakamura, H. Miyashita, J. Minamino, S. Furumiya, T. Ishida, and H. Ishibashi, *Jpn. J. Appl. Phys., Part 1* 41, 1687 (2002).
2. T. Ohta, K. Nishiuchi, K. Narumi, Y. Kitaoka, H. Ishibashi, N. Yamada, and T. Kozaki, *Jpn. J. Appl. Phys.* 39, 770 (2000).
3. A. Partovi, D. Peale, M. Wutting, C. Murry, G. Zydzik, L. Hopkins, K. Baldwin, W. Hobson, J. Wynn, J. Lopata, L. Dhar, R. Chichester, and J. Yeh, *Appl. Phys. Lett.* 75, 1515 (1999).
4. K. Hirota, T. Milster, Y. Zhang, and J. Erwin, *Jpn. J. Appl. Phys.* 39, 973 (2000).
5. W. Hung and M. Mansuripur, *Appl. Opt.* 39, 302 (2000).
6. F. Guo, T. Schlesinger, and D. Stancil, *Appl. Opt.* 39, 324 (2000).
7. S. Jiang, H. Ohsawa, K. Yamada, T. Pagaribuan, M. Ohtsu, K. Imai, and A. Ikai, *Jpn. J. Appl. Phys.* 31, 2282 (1992).
8. M. Lee, M. Kourogi, T. Yatsui, K. Tsutsui, N. Atoda, and M. Ohtsu, *Appl. Opt.* 38, 3566 (1999).
9. J. Tominaga, H. Fuji, A. Sato, T. Nakano, and N. Atoda, *Jpn. J. Appl. Phys.* 39, 957 (2000).
10. H. Fuji, J. Tominaga, L. Men, T. Nakano, H. Katayama, and N. Atoda, *Jpn. J. Appl. Phys.* 39, 980 (2000).
11. K. Itoh, H. Yoshimura, and K. Ogawa, *Jpn. J. Appl. Phys.* 39, 714 (2000).
12. H. Awano, M. Sekine, M. Tani, N. Kasajima, N. Ohta, K. Mitani, N. Takagi, and S. Sumi, *Jpn. J. Appl. Phys.* 39, 725 (2000).
13. J. H. Kim, D. Buechel, T. Nakano, J. Tominaga, N. Atoda, H. Fuji, and Y. Yamagawa, *Appl. Phys. Lett.* 77, 1774 (2000).
14. J. Tominaga, D. Buechel, T. Nakano, H. Fuji, T. Fukaya, and N. Atoda, *Proc. SPIE* 4081, 86 (2000).
15. H. Awano, H. Shirai, N. Ohta, A. Yamaguchi, S. Sumi, and K. Torazawa, *Nippon Ouyoh Jiki Gakkaishi* 22, 337 (1998) (in Japanese).
16. H. Raether, "Surface Plasmons." Springer, Berlin, 1988.
17. V. M. Shalaev, "Nonlinear Optics of Random Media." Springer, Berlin, 2000.
18. M. Mansuripur, "The Physical Principles of Magneto-Optical Recording," p. 638. Cambridge University Press, Cambridge, 1995.
19. M. Okuda, H. Naito, and T. Matsushida, *Jpn. J. Appl. Phys., Part 1* 31, 466 (1992).
20. J. Tominaga, T. Nakano, and N. Atoda, *Appl. Phys. Lett.* 73, 2078 (1998).
21. H. Fuji, J. Tominaga, L. Men, T. Nakano, H. Katayama, and N. Atoda, *Jpn. J. Appl. Phys., Part 1* 39, 980 (2000).
22. J. H. Kim, M. Kuwahara, N. Atoda, and J. Tominaga, *Appl. Phys. Lett.* 79, 2600 (2001).
23. S. R. Lee and A. E. Miller, *J. Appl. Phys.* 55, 3465 (1984).
24. X. S. Miao, T. C. Chong, L. P. Shi, P. K. Tan, J. M. Li, and K. G. Lim, *Jpn. J. Appl. Phys., Part 1* 40, 1581 (2001).
25. N. Yamada, M. Otoba, K. Kawahara, N. Miyagawa, and H. Ohta, *Jpn. J. Appl. Phys., Part 1* 37, 2104 (1998).
26. R. B. Van Dover, E. M. Gyorgy, R. P. Frankenthal, M. Hong, and D. J. Siconolfi, *J. Appl. Phys.* 59, 1291 (1986).
27. F. Luborsky, J. T. Furey, R. E. Skoda, and B. C. Wagner, *IEEE Trans. Magn.* MAG-22, 937 (1986).
28. T. Anthony, J. Brug, S. Naberhuis, and J. Birecki, *Appl. Phys.* 59, 213 (1985).
29. H. E. Kissinger, *Anal. Chem.* 29, 1702 (1959).
30. L. Men, J. Tominaga, H. Fuji, T. Kikukawa, and N. Atoda, *Jpn. J. Appl. Phys., Part 1* 40, 1629 (2001).
31. Yasui, "Optical Materials," p. 24. Dainippon, Tokyo, 1991 (in Japanese).
32. C. Joachim, J. K. Gimzewski, and A. Aviram, *Nature* 408, 541 (2000).
33. M. A. Reed, C. J. Zhou, T. P. Burgin, and J. M. Tour, *Science* 278, 252 (1997).
34. C. Kerguelis, *Phys. Rev. B* 59, 12505 (1999).
35. A. Bezryadin, C. Dekker, and G. Schmid, *Appl. Phys. Lett.* 71, 1273 (1997).
36. E. Di Fabrizio, L. Grella, M. Gentili, M. Baciocchi, L. Mastrogiacomo, and P. Morales, *Jpn. J. Appl. Phys.* 36, L70 (1997).
37. J. H. Kim, H. Akinaga, and J. Tominaga, *Appl. Phys. Lett.* 80, 2764 (2002).
38. S. M. Sze, "Semiconductor Devices," p. 295. Wiley, New York, 1985.
39. C. Hu and R. M. White, "Solar Cell," p. 55. McGraw-Hill, New York, 1983.
40. C. Kittel, "Introduction to Solid State Physics," p. 148. Wiley, New York, 1986.
41. Y. Kasami, K. Yasuda, M. Ono, A. Fukumoto, and M. Kenako, *Jpn. J. Appl. Phys., Part 1* 35, 423 (1996).
42. E. Betzig, J. K. Trautman, R. Wolfe, E. M. Gyorgy, P. L. Finn, M. H. Kryder, and C. H. Chang, *Appl. Phys. Lett.* 61, 142 (1992).
43. T. J. Silva and S. Schultz, in "Near Field Optics" (D. W. Pohl and D. Courjon, Eds.), p. 263. Kluwer Academic, Arc-et-Senans, France, 1993.
44. T. J. Silva and S. Schultz, *Appl. Phys. Lett.* 65, 658 (1994).
45. A. P. Michael and J. M. Patrick, "Near-Field Optics," p. 67. Wiley, New York, 1996.
46. E. Betzig, S. G. Grubb, R. J. Chichester, D. J. DiGiovanni, and J. S. Weiner, *Appl. Phys. Lett.* 63, 3550 (1993).
47. B. D. Terris, H. J. Mamin, D. Ruger, W. R. Studenmund, and G. S. Kino, *Appl. Phys. Lett.* 65, 388 (1994).
48. H. J. Mamin, B. D. Terris, and D. Ruger, "Proceedings of MORIS," 1994.
49. K. Kishima, I. Ichimura, K. Saito, K. Yamamoto, Y. Kuroda, A. Iida, S. Masuhara, and K. Osato, *Jpn. J. Appl. Phys., Part 1* 41, 1894 (2002).
50. I. Ichimura, K. Kishima, K. Osato, K. Yamamoto, Y. Kuroda, and K. Saito, *Jpn. J. Appl. Phys., Part 1* 39, 962 (2000).
51. H. Awano, S. Ohnuki, H. Shirai, N. Ohta, A. Yamakuchi, S. Sumi, and K. Torazawa, *Appl. Phys. Lett.* 69, 4257 (1996).
52. H. Awano, M. Sekine, M. Tani, N. Kasajima, N. Ohta, K. Mitani, N. Takagi, and S. Sumi, *Jpn. J. Appl. Phys., Part 1* 39, 725 (2000).
53. H. Awano, M. Sekine, M. Tani, N. Kasajima, N. Ohta, K. Mitani, N. Takagi, and S. Sumi, *Proc. SPIE* 3864, 66 (1999).
54. S. H. Wang and A. M. Taratorin, in "Magnetic Information Storage Technology," Chap. 16. Academic, San Diego, 1999.
55. W. Challener, T. McDaniel, C. Mihalcea, and K. Sendur, "ISOM/ODS 2002," 2002, p. 159.
56. X. Shi, R. L. Thornton, and L. Hesselink, *Proc. SPIE* 4342, 320 (2002).
57. A. Itagi, F. Chen, D. D. Stancil, and T. E. Schlesinger, *Proc. SPIE* 4342, 277 (2002).
58. M. Alex, A. Tselikov, T. McDaniel, N. Deeman, T. Valet, and D. Chen, *IEEE Trans. Magn.* 37, 1244 (2001).
59. H. Sukeda, H. Saga, H. Nemoto, Y. Itou, C. Haginoya, and T. Matsumoto, *IEEE Trans. Magn.* 37, 1234 (2001).
60. K. Ito, H. Nemoto, C. Haginoya, Y. Kawato, and H. Sukeda, "ODS 2000," 2000, p. 30, paper MC-1.
61. K. Shimazawa, O. Redron, N. Kasahara, J. Sun, S. Araki, H. Morota, and M. Matsuzaki, "Intermag 2000," 2000, paper FA-01.

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanostructured Carbide-Derived Carbon

A. Nikitin, Y. Gogotsi

Drexel University, Philadelphia, Pennsylvania, USA

CONTENTS

1. INTRODUCTION

Carbon is a fundamental building block of material in living organisms. Carbon is present in all organic substances and a whole class of materials—polymers—is based on carbon compounds. Pure elemental carbon has been one of the most important inorganic materials for a long time. Carbon is the only element that has a major monthly scientific journal named after it which publishes papers on carbon materials only (*Carbon*, published by Elsevier Science). One of the reasons for this is that carbon exists in four different allotropes—graphite, diamond, fullerene, and carbine—which have numerous applications. Moreover, diamond and fullerenes have journals dedicated to them (*Diamond and Related Materials* and *Fullerene Science and Technology*). The recently discovered carbon nanotubes are the most important building blocks in nanotechnology and are considered to be the strongest fibers known. It is amazing that carbon materials can cover the whole range of properties that scientists and engineers need. For example, diamond is the hardest material known (microhardness of above 100 GPa) and is used to make cutting tools, while graphite is one of the softest materials (microhardness is about 1 GPa) and is used as a solid lubricant. While diamond is a wide-band-gap semiconductor, graphite is a good electrical conductor. Carbon nanotubes can have the conductivity of copper (metallic single-wall nanotubes) or silicon (semiconducting single wall nanotubes). In addition to the known carbon allotropes, amorphous or disordered carbons form a whole new class of materials for which all properties, such as electrical conductivity, hardness, wear resistance, gas permeability, optical transparency, and many others, can be finely tuned by controlling the sp^2/sp^3 carbon ratio, porosity/density, size and ordering of nanocrystals, and inclusions of various allotropic forms of carbon and other elements.

There are many ways to produce carbon materials. Very high temperatures are typically required for the synthesis of ordered graphite because of the very low mobility of carbon atoms in its covalently bonded layers. Diamond is produced either at high pressures or by using plasma and other high-energy activation techniques. Fullerenes and nanotubes are manufactured in plasma as well. Chemical vapor deposition (CVD) is currently a common method for synthesis of thin and thick films of diamond, disordered graphitic carbon, and nanotubes. Many porous and disordered carbons, including carbon fibers and activated carbons, are made by thermal decomposition of organic compounds. The age of nanotechnology requires new methods that enable engineers to control the growth of carbon materials at the atomic level, provide better structural control, and have the capability to produce novel carbon structures such as nanotubes, onions, nanocrystalline diamond or nanoporous carbon with a narrow pore size distribution for gas storage, batteries, supercapacitors, and other applications.

It has previously been shown that selective etching of carbides is an attractive technique for the synthesis of various carbon structures [1]. Carbon produced by extraction of metals from carbides is called carbide derived carbon (CDC). Leaching in supercritical water, high-temperature treatment in halogens, vacuum decomposition, and other methods can be used to remove metals from carbides to produce carbon coatings or bulk and powdered carbon. Linear reaction kinetics has been demonstrated for the chlorination of SiC [2] which allowed transformation to any depth up to the complete conversion of the particle or component to carbon. CDC demonstrated a variety of carbon phases depending on experimental conditions. Amorphous and nanocrystalline graphitic carbon [3], nanotubes [4], carbon onions, nanocrystalline diamond [5], and ordered graphite have been reported. Thus, this method allows the synthesis of almost every carbon allotrope. Since the carbide lattice is used as a template and metal is extracted layer by layer, atomic-level control can be achieved in the synthesis

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (553–574)

process. The structure of the carbon can be templated by the carbide structure, with the opportunity for further structural modification by controlling the temperature, environment, and other process variables. An important advantage of the CDC process over CVD or physical vapor deposition (PVD) for coating synthesis is that the transformation is conformal and does not lead to change in the carbide sample size and shape. This may be of particular importance in nanotechnology, when nanometer range accuracy must be achieved.

Significant progress in CDC synthesis has been achieved over the past decade, but no review of the carbon synthesis by extraction of metals from carbides has been published. This chapter presents a summary of the work done in this exciting area that shows various approaches to CDC synthesis (leaching by supercritical water, halogens, salts, and decomposition in vacuum), analyzes similarities and differences between different techniques, and describes the variety of structures produced, their applications, and the potential of this method for producing new carbon structures with useful properties.

2. SELECTIVE ETCHING OF CARBIDES BY HALOGENS

2.1. Chlorination of Carbides for Production of Chlorides

A manufacturing process based on the reaction of silicon carbide with chlorine gas was reported in 1918 for the first time [6]. Hutchins patented a method for the production of silicon tetrachloride using the reaction

$$SiC + 2Cl_2 = SiCl_4 + C \quad (1)$$

The reaction was conducted by passing dry chlorine over silicon carbide heated to a temperature of 1000 °C or above. For this purpose a resistance furnace, illustrated in Figure 1, was developed consisting of a vertical cylindrical shell with a carbon heating rod inside it. The main product of the reaction, silicon tetrachloride vapor, passed through the outlet pipe into a condenser. Residual carbon, considered to be "in the form of graphite," was disposed. An improvement of this method was patented in 1956 [7]. In order to remove carbon without cooling off the reactor, air was periodically blown through it to remove carbon in the form of gaseous oxides CO and CO_2. This method has been used for the commercial production of $SiCl_4$.

Treatment of other metal carbides in the chlorine gas was used for the purpose of producing corresponding metal chlorides. Therefore, the kinetics of chlorination for titanium and niobium carbides [8] and zirconium carbide [9] was studied. The reactions

$$MeC + x/2Cl_2 = MeCl_x + C \quad (2)$$

where $x = 4$ when Me = Ti, Zr and $x = 5$ when Me = Nb, were conducted in a quartz tube flow reactor in a temperature range of 600–800 °C in the case of TiC and NbC and 450–950 °C in the case of ZrC. Samples were produced by hot-pressing the carbide powders. Reaction rates

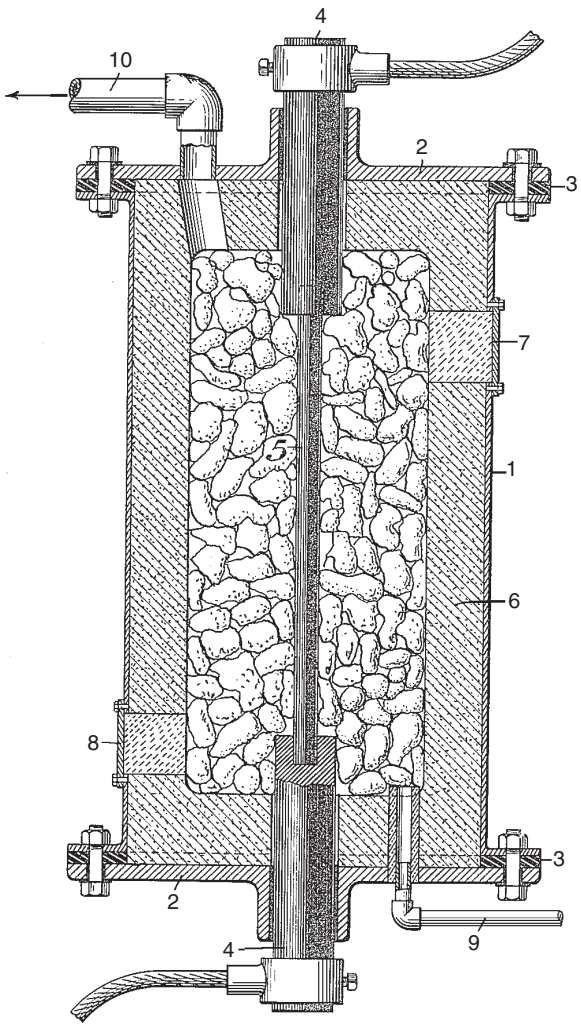


Figure 1. Vertical furnace for the silicon tetrachloride production [6]. (1) Steel shell; (2) head; (3) insulating gasket; (4) carbon terminal; (5) carbon resistor rod; (6) shell lining; (7) opening; (8) opening; (9) inlet pipe; (10) outlet pipe.

were monitored by measuring the weight loss and the amount of formed chlorides. Carbon formed on the surface of the samples in the course of the reaction was considered to be porous graphite or a carbonaceous "cinder." This layer did not affect the reaction rate at 400 °C but at higher temperatures the diffusion resistance was considerable. In the case of niobium carbide, the chlorination rate at 600 °C was higher than at 800 °C. This phenomenon was initially attributed to increased adhesion of the carbon layer to the carbide surface, but our recent work shows that it should primarily be attributed to changes in the CDC structure. For some carbides, the mechanism of the interaction with chlorine and the composition of reaction products may change with temperature. Therefore, it is important to understand the thermodynamics of the reactions of carbides with halogens.

2.2. Thermodynamic Simulations

Thermodynamic simulations were performed to predict the equilibrium composition of the chlorination products for a number of carbides [10, 11]. Calculations were done using

Gibbs energy minimization software in the temperature range from 300 to 1300 °C. The initial systems included carbide and chlorine gas with optional addition of hydrogen and oxygen at various pressures and carbide/gas molar ratios.

The calculations were done for a closed system with constant total pressure: the solid phase was treated as a mechanical mixture of compounds with unit activities and the gas phase as a mixture of ideal gases. It was found that carbon formation is expected as a result of halogenation of SiC, TiC, ZrC, V_2C, NbC, Nb_2C, TaC, Ta_2C, Cr_3C_2, Mo_2C, MoC, W_2C, WC, BaC_2, CaC_2, SrC_2, Al_4C_3, B_4C, and Fe_3C (see Fig. 2). In the SiC–Cl_2(g) system, reaction (1) is most favorable. At higher temperatures (above 900 °C) probability of $SiCl_3$(g) formation increases. In the SiC–Cl_2(g)–H_2(g) system the predicted amount of carbon is somewhat lower than that in the SiC–Cl_2(g) system and it decreases with increasing H_2:Cl_2 ratio. The maximum equilibrium amount of carbon is obtained around 800 °C (see Fig. 2a). At increased pressure, the temperature at which the carbon content in the reaction products is maximized increases. A small addition of oxygen leads to the formation of silica but does not affect the relative amounts of other species in the system. The calculations are generally in good agreement with experimental results. We have conducted thermodynamic simulations for interactions of a variety of carbides with chlorine and other halogens as well as halogen-containing gases such as HCl and HF. It may be concluded that a very large number of carbides can be used for CDC synthesis.

2.3. Effect of Halogenation Techniques on Properties of CDC

CDC did not attract the attention of researchers before 1959, when Mohun introduced "a fourth class of amorphous carbons" along with the three known ones (hard carbons, soft carbons, and carbon black) [3]. The term "mineral carbons" was proposed for carbons formed by chlorination of inorganic carbides, whose corresponding chlorides are volatile, followed by elimination of the residual chlorine, to distinguish them from carbons of organic origin. Unlike the latter, whose composition also includes substantial quantities of hydrogen and other elements, CDC was claimed to be real elemental amorphous carbon. Chlorination of silicon, aluminum, boron, titanium, and zirconium carbides as well as zirconium carbonitride was conducted in the temperature range from 150 °C for ZrC · ZrN to 1650 °C for SiC. Reactor tubes were made of quartz or mullite and provided with an external electric heater. Due to the strong exothermic effect of the reaction and the substantial size of the samples (20–250 g) the temperature of the sample was rather difficult to control. After the chlorination process was finished, the carbon samples were cooled in an atmosphere of chlorine or dry nitrogen. Then the samples cooled in the chlorine atmosphere were treated to remove chlorine.

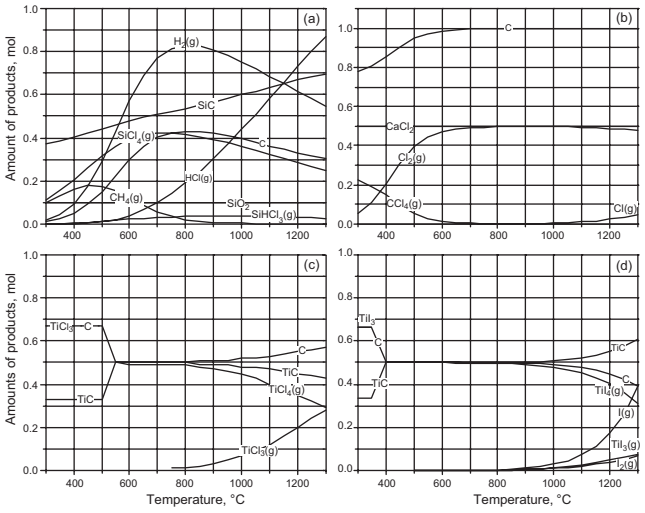


Figure 2. Examples of calculated equilibrium composition for the carbide/halogen systems at ambient pressure. Initial compositions: (a) Cl_2 (0.9 mol)–H_2 (1 mol)–O_2 (0.1 mol)–SiC (1 mol); (b) Cl_2 (1 mol)–CaC_2 (0.5 mol); (c) Cl_2 (1 mol)–TiC (1 mol); (d) I_2 (1 mol)–TiC (1 mol).

Mohun extended his study and eventually patented the new materials and the process of their production [12, 13]. Methods employed for the dechlorination of the CDC samples were described in detail and included heating in the static air, heating in a gas stream (air, nitrogen, water vapor, and carbon dioxide), vacuum dechlorination, and wet dechlorination with the use of various washing solutions. Chlorine content in the samples was monitored by measuring weight loss during dechlorination and the pH of the aqueous suspension. It was found that passing nitrogen or another gas over the sample heated to 150 °C at atmospheric pressure provided rapid and complete dechlorination. Of considerable importance are the chlorides which remain trapped in the carbon when the reaction is complete. It was shown that they are readily removed by any cyclic fluctuation in temperature or pressure carried out at the same time as dechlorination. Some physical and chemical properties of the CDC reported in these patents are summarized in Table 1. Another interesting property of this CDC was the absence of measurable paramagnetic electron spin absorption. This distinguishes CDC from organic chars in which the free radicals are trapped during the carbonization of organic matter.

The chlorination setup used in this study is a common setup for heat treatment of solid samples in the flowing gases. A version of the apparatus is presented in Figure 3.

2.4. Nanoporous Structure and Adsorption Properties

The porous structure and molecular sieve properties of CDC were first investigated by Boehm and Warnecke in 1975 [14]. Carbon samples were prepared by treatment of tantalum and silicon carbides in flowing chlorine at 500 and 900 °C respectively. Completeness of the reaction was confirmed by the disappearance of the X-ray diffraction (XRD) reflexes of the original carbides. The samples from TaC were then heat-treated at various temperatures.

The adsorption isotherms of CDC were measured for N_2 at −196 °C, for CO_2 at −79 and 0 °C, and for a variety of organic adsorptives at 25 °C. All the isotherms were of type I in the Brunauer classification, which indicates a nanoporous structure in the CDC samples under study. There is a maximum in the N_2 uptake for CDC from TaC after

Table 1. Properties of CDC from SiC [12, 13].

	Temperature of chlorination (°C)	1050	1200	1650	1050	825
	Atmosphere during cooling	Cl_2	Cl_2	Cl_2	N_2	N_2
	Temperature of static dechlorination in the air (°C)	120	120	110	—	—
Mesh	Electrical resistivity (Ω cm)	0.044	0.042	—	0.035	0.51
150	Vapor adsorption capacity (g/g)					
	m-Xylene	0.377	0.380	0.204	0.515	0.316
	Acetone	0.321	0.330	0.183	0.493	0.320
	Benzene	0.403	0.405	0.318	0.572	0.471
	Water	0.169	0.159	0.153	0.155	0.238
	Heat of wetting (kJ/g)					
	m-Xylene	45.2	48.5	20.5	71.5	46.4
	Acetone	46.4	48.5	24.2	64.9	33.1
	Benzene	44.3	46.9	20.5	58.6	40.6
	Water	27.2	13.4	9.2	24.3	37.2
Mesh	Adsorption from solution (g/g)					
325	Iodine	1.230	1.340	0.647	1.100	1.435
	Phenol	0.344	0.368	0.249	0.318	0.296
	Methylene blue	0.336	0.272	0.330	0.336	0.329
	pH of aqueous suspension	2.4	2.5	3.2	7.1	5.2

heat treatment at 1450 °C. The low capacity of the carbon prepared at 500 °C was explained by the presence of tantalum pentachloride, which blocked a considerable part of the pore volume. However, our experience shows that the very small pore size can limit applicability of the sorption techniques for analysis of CDC produced from transition metals carbides at ~500 °C. The maximum pore volume accessible to N_2 in CDC after heat treatment at 1450 °C was 0.644 cm^3/g and this is in excellent agreement with the theoretically calculated value of 0.645 cm^3/g. Evaluation of the CO_2 isotherms of the CDCs heat-treated in the range 500–1100 °C gave a maximum pore volume of 0.273 cm^3/g.

Adsorption of organic vapors (benzene, tetrachloromethane, toluene, and α-pinene) demonstrated strong molecular sieve effects for CDC from TaC heat-treated above 1500 °C. It was possible to separate low molecular weight gases using CDC from SiC treated in hydrogen at 800 °C, proving it to be suitable for gas chromatography. For example, O_2, N_2, CO, and CO_2 could be separated with a 1 m column at 30 °C; CH_4, C_2H_2, C_2H_4, and C_2H_6 were separable at 100 °C.

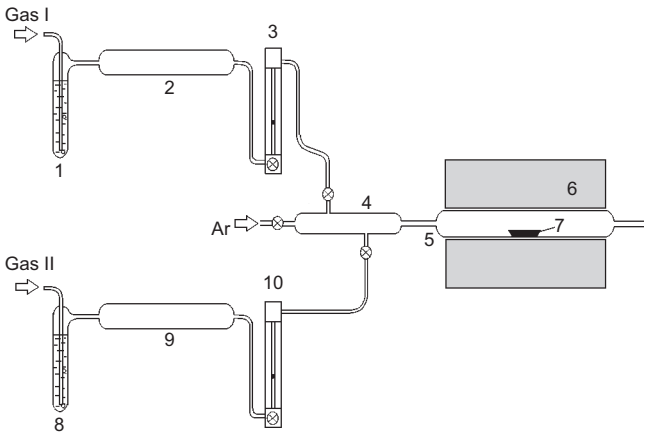


Figure 3. Schematic of a laboratory chlorination setup. (1, 8) Sulfuric acid; (2, 9) desiccant; (3, 10) flowmeter; (4) mixing column; (5) fused silica reaction tube (3–8 cm in diameter); (6) resistance furnace; (7) sample.

The porous structure and adsorption properties of CDC from silicon carbide and carbides of the transition metals of IV–VI groups were systematically studied by Fedorov et al. in the early 1980s [15–18]. Three forms of silicon carbide (green and black α-SiC and β-SiC, no temperature indicated) [15]; titanium carbide in the temperature range 500–1100 °C [16]; carbides of zirconium, hafnium, vanadium, niobium, tantalum, chromium, and molybdenum at 700 °C; and tungsten monocarbide, WC, at 1100 °C [17] were treated in flowing chlorine in a quartz tube reactor. Vapors of bromine and iodine were used in addition to chlorine for etching zirconium carbide [18]. Completeness of the reaction was determined by the measuring weight loss during the process and by the amount of carbide in the sample.

The porous structure of the obtained CDCs was studied by a combination of pycnometric, mercury-porosimetric, and sorption methods. Theoretical values of the apparent density δ were calculated (assuming that no shrinkage of the samples occurred during the chlorination process, which was confirmed by electron microscopy) using the formula

$$\delta = dA \tag{3}$$

where d is the carbide density and A is the mass fraction of carbon in the carbide. Theoretical total pore volume V_Σ was calculated on the basis of the given pycnometric density of carbon ($d_C = 2.00$ g/cm^3) by the formula

$$V_\Sigma = \frac{1}{\delta} - \frac{1}{d_C} \tag{4}$$

It was found that the theoretical δ and V_Σ values are generally in satisfactory agreement with the experimental data, provided the porosity of the initial carbide granules is accounted for.

According to the IUPAC definition, porosity in solids is classified by pore size [19]:

(i) pores with widths exceeding about 50 nm are called *macropores*;
(ii) pores of widths between 2 and 50 nm are called *mesopores*;
(iii) pores with widths not exceeding about 2 nm are called *micropores*.

We use the designation *nanopores* for the pores of ≤2 nm because this term reflects their actual size more properly.

Sorption analysis was carried out with the use of benzene vapor as an adsorptive. The isotherms were of type I in the Brunauer classification, characterizing these carbons as nanoporous adsorbents. The sole exception was CDC from molybdenum carbide, which had developed mesopore volume. This was explained as a result of the atomic ratio between metal and carbon which is twice that of other transition metal carbides. The isotherms were evaluated using Dubinin's theory of the volume filling of nanopores. The characteristic energy of adsorption, which is the principal parameter used to describe nanoporous structure, of the CDCs in question was comparable to that of the best activated carbons.

The temperature dependence of the porous structure of CDC from TiC is of particular interest [16]. It was shown that the sample produced at 500 °C had a higher apparent density, which was explained by presence of titanium tetrachloride's occupation of a significant part of the pore volume. Virtually equal values of total pore volume were observed for the samples produced at 700, 900, and 1100 °C while the pore size distributions were different; the mesopore volume increased with increasing chlorination temperature as a result of three-dimensional ordering of turbostratic carbon structures. The mesopore volume measured by benzene adsorption and mercury porosimetry (pores of equivalent radius 3–200 nm) is presented in Figure 4. This is evidence that CDC produced at 1100 °C had a part of mesopores of equivalent radius less than 3 nm.

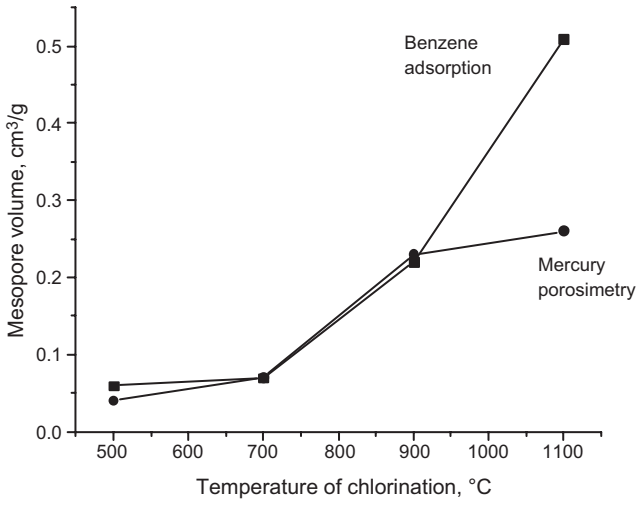


Figure 4. Mesopore volume of CDC from TiC [16].

The temperature dependence of the porous structure of CDC produced from TiC in the presence of 3d-transition-metal catalysts was obtained in [20]. Titanium carbide samples were chlorinated for 1 hour over the wide temperature range of 400–1200 °C. Pure TiC samples as well as TiC with the addition of a mixture of cobalt and nickel dichloride and ferric chloride were treated. The specific surface area, S_{BET}, was calculated using BET theory from the data of nitrogen sorption at −196 °C; the nanopore volume W_{na} was determined by measuring dynamic benzene adsorption at room temperature (see Table 2).

The chlorination of aluminum carbide and the properties of carbon materials produced thereby were reported in [21, 22]. The nitrogen sorption isotherm of a low temperature carbon (300 °C) was of a type I in the Brunauer classification, indicating this sample has a nanoporous structure. The isotherms of materials produced at 500–900 °C were of type IV, which indicates the presence of mesopores [19]. Total pore volume (V_Σ), average pore size according to BJH theory, specific surface area according to BET theory, and nanopore volume using *t* plots were calculated. The results are summarized in Table 3.

The porous structure of CDC from ZrC depends on the nature of the reacting halogens as was shown in [18]. Equal values of total pore volume consisting purely of nanopores were found by measuring benzene adsorption in the samples prepared by treatment in chlorine, bromine, and iodine at 800 °C. At the same time their sorption capacities for halogens differed significantly. The highest capacities for bromine and iodine adsorption were observed to correspond to CDC samples prepared by bromination and iodination respectively. It was considered as evidence that the porous structure of CDC was "imprinted" by the size of the halogen and the evolving zirconium tetrahalogenide molecules.

Fedorov summarized his experimental results [15–17] with the addition of previously unpublished data in 1995 [23] (see Table 4). CDC samples from boron and uranium carbides and single crystals of titanium, zirconium, and niobium carbides were produced by chlorination at 700 °C. It is notable that CDC from single crystals absolutely did not have macroporosity, distinguishing it from carbons produced from sintered and powder samples of carbides having initial porosity of 6 to 16% vol. This means that the original macropore volume of sintered carbides did not change during the chlorination process and can be controlled by controlling the

Table 2. Properties of carbon synthesized from TiC with no catalyst (NOC) and with catalysts (CAT).

	S_{BET} (m^2/g)		V_{na} (cm^3/g)		q (%)	
T_{Chlor} (°C)	NOC	CAT	NOC	CAT	NOC	CAT
400	1180	1370	0.441	0.550	0	0
500	1350	1441	0.510	0.607	0	0
600	1407	1390	0.495	0.543	0	0
700	1400	1470	0.508	0.621	0	0
800	1580	1200	0.600	0.497	0	12
900	1350	876	0.638	0.324	5	28
1000	1563	182	0.682	0.097	6	35
1100	1440	193	0.65	0.116	11	37
1200	1070	138	0.696	0.111	16	47

Table 3. Properties of carbon synthesized from Al_4C_3.

T_{Chlor} (°C)	S_{BET} (m^2/g)	V_{na} (cm^3/g)	V_{Σ} (cm^3/g)	D (nm)	q (%)
300	1353	0.507	0.641	1.43	0
500	1224	0.304	0.799	1.72	29.0
700	716	0.065	0.836	3.09	41.6
900	683	0.038	0.949	3.57	46.0

porosity of the carbides. The total volume and characteristic dimensions of meso- and nanopores can be predicted and achieved by the selection of proper binary or ternary carbide and variation of the chlorination process parameters. This statement is illustrated with the example of CDC produced from a number of TiC–HfC solid solutions of variable composition. The linear dependence of the total pore volume and characteristic energy of benzene adsorption on composition is shown in Figure 5.

A systematic study of the significance of the initial carbide porosity on the porosity of the corresponding CDC was reported [24, 25]. Four dense pyrolitic carbides produced by CVD from $MeCl_4$(g)–CH_4(g)–H_2(g) (Me = Si, Ti, Hf, and Zr) were treated in chlorine gas at 600–1100 °C [24]. As in the case of CDC from single-crystal carbides, the CDC materials from pyrolitic carbides were completely nanoporous, and the experimental values of apparent density and total pore volume were in good agreement with theoretical ones. The data on the porous structure of these materials are also presented in Table 4. The porous structure of the carbon materials prepared by chlorination of silicon carbide with various initial porosities was investigated in the work [25] and is presented in Table 5. The authors concluded that the uniformity of the pore size and shape makes CDC a good model system for the investigation of the sorption process in nanopores. Eventually, these CDCs were used for a study of the association of water molecules in nanopores [26, 27].

The average pore sizes for some materials described in [23] were calculated in the review by Gordeev [28] and data on specific surface area are given in Table 6.

The calculation of the porosity of CDC in the general case for the reaction

$$Me_kC_l + kp/2nCl_2 = k/nMe_nCl_p + lC \qquad (5)$$

was carried out in the theoretical work [29]. It was found that the total pore volume per unit volume is equal to

$$V_V = d(1-\varepsilon_0)\left(\frac{1}{d}\frac{\varepsilon_0}{1-\varepsilon_0} + \frac{\alpha}{d} - \frac{\alpha}{d_C}\frac{lM_C}{M}\right) \qquad (6)$$

where V_V is the total pore volume per unit volume (cm^3/cm^3), ε_0 is the total pore volume of the initial carbide, α is the fractional conversion of reaction (6), and M_C and M are the molecular masses of the carbon and carbide.

A novel method for the production of nanoporous carbon articles of any shape having sufficient mechanical strength was developed by Gordeev et al. [30–32] and Fedorov et al. [33]. In this method, an intermediate blank of metal carbide having a porosity in the range of 30% to 50% by volume is formed in the desired shape. The blank is then impregnated with carbon as a binder by exposing it to a gaseous

Table 4. Parameters of porous structure of CDC [15–17, 23, 24].

Initial carbide	Parameters of porous structure										
	Calculated			Experimental							
	d (g/cm^3)	δ (g/cm^3)	V_{Σ} (cm^3/g)	δ (g/cm^3)	ρ (g/cm^3)	V_{Σ} (cm^3/g)	V_{ma} (cm^3/g)	V_{me} (cm^3/g)	V_{na} (cm^3/g)	W_0 (cm^3/g)	E_0 (kJ/mol)
α-SiC (g)	3.21	0.96	0.52	0.99	2.06	0.52	0.03	0.05	0.44	0.45	23.1
α-SiC (b)				0.94	2.02	0.53	0.03	0.03	0.47	0.48	22.4
β-SiC				0.97	1.98	0.64	0.04	0.04	0.43	0.45	21.7
SiC[a]				0.94	2.24	0.62	0.00	0.00	0.62	0.61	
TiC	4.93	0.99	0.50	0.87	2.10	0.67	0.06	0.07	0.54	0.54	23.5
TiC[b]				0.98	2.09	0.54	0.00	0.04	0.50	0.51	24.5
TiC[a]				0.99	2.14	0.54	0.00	0.01	0.53	0.53	
ZrC	6.70	0.78	0.61	0.77	2.20	0.85	0.11	0.09	0.65	0.66	22.1
ZrC[b]				0.85	2.14	0.71	0.00	0.03	0.69	0.69	21.8
ZrC[a]				0.78	2.17	0.82	0.00	0.01	0.81	0.81	
HfC	12.2	0.77	0.80	0.78	2.15	0.82	0.09	0.07	0.66	0.63	18.0
HfC[a]				0.78	2.04	0.79	0.00	0.01	0.78	0.78	
NbC	7.82	0.90	0.55	0.77	2.02	0.80	0.15	0.12	0.53	0.55	22.3
NbC[b]				0.93	2.10	0.60	0.00	0.03	0.57	0.57	22.2
Mo_2C	8.90	0.53	0.74	0.49	1.08	1.21	0.10	0.85	0.26	0.27	15.1
Cr_2C_3	6.68	0.89	0.55								
WC	15.63	0.95	0.55	0.86	2.10	0.69	0.11	0.09	0.49	0.50	21.5
TaC	14.50	0.90	0.61	0.84	2.15	0.73	0.12	0.11	0.50	0.51	21.9
UC	5.46	1.04	0.48	0.90	2.08	0.63	0.13	0.06	0.44	0.47	24.3
B_4C	2.52	0.55	0.72	0.72	2.18	0.67	0.13	0.10	0.44		

Note: d is carbide density; δ is apparent density; V_{Σ} is total pore volume; ρ is pycnometric density; V_{ma}, V_{me}, V_{na} are macro-, meso-, and nanopore volumes; W_0 is maximum adsorption volume; E_0 is characteristic energy of adsorption.

[a] Nonporous polycrystalline material.

[b] single crystal.

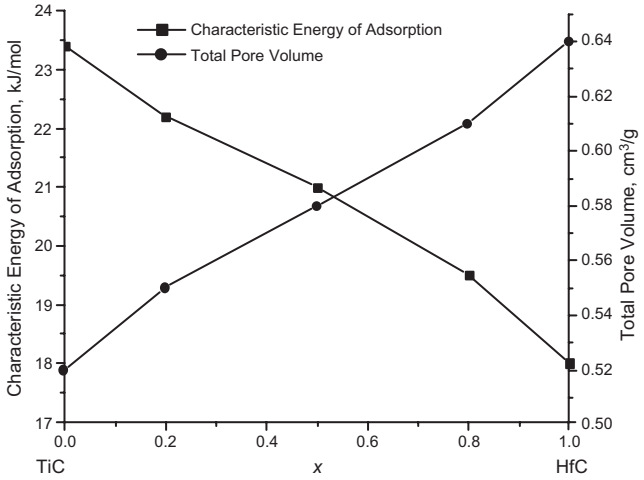


Figure 5. Characteristic energy of adsorption and total pore volume of CDC produced from solid solutions of $Ti_{1-x}Hf_xC$ [23].

hydrocarbon mixture at a temperature exceeding the decomposition temperature of hydrocarbons (700–1000 °C) until the mass of the blank increases by 10%–25% due to the formation of pyrocarbon. The blank is then thermochemically (900–1100 °C) treated in gaseous chlorine. A nanoporous structure is formed during removal of volatile chlorides of the carbide-forming elements according to reaction (5). In a variation of the method, the silicon carbide blank may be impregnated with liquid silicon, with which the hydrocarbon mixture interacts to form a secondary silicon carbide. Removal of the free and bonded silicon in the form of silicon tetrachloride occurs during the following thermochemical treatment. The porous structure of the prepared carbon articles was studied by the hydrostatic weighing method to determine total pore volume and apparent density, and by measurement of benzene adsorption in saturation conditions to determine the nanopore volume. The specific surface area was determined by evaluating nitrogen adsorption isotherms at −196 °C according to the BET and Langmuir theories. The obtained data are presented in Table 7. Typical pore size distribution in a composite material prepared using TiC is illustrated in Figure 6 [32].

It was shown that macroporosity developed in the course of the formation of the carbide–pyrocarbon composite remained unchanged after thermochemical treatment in the chlorine environment [33]. Thus, adsorbents with a controlled diporous structure were produced. The influence of pyrocarbon content on the porous structure and physical properties, including resistivity, heat conductivity, and compression strength, was systematically studied [34] for

Table 5. Parameters of porous structure of CDC materials from SiC with various initial porosities and their compression strength.

Initial SiC porosity (%)	δ (g/cm^3)	ρ (g/cm^3)	V_Σ (cm^3/cm^3)	V_{na} (cm^3/cm^3)	Compression strength (MPa)
20	0.80	2.25	0.64	0.42	6–7
40	0.65	2.24	0.71	0.37	1.8–4.0
48	0.48	2.22	0.78	0.27	failure

Table 6. Pore size and specific surface area (S_{BET}) of CDC.

Initial carbide	SiC	TiC	ZrC	NbC	Mo_2C	B_4C
Pore size (nm)	0.8	0.8	1.2	1.2	2.1	2.0
S_{BET} (m^2/g)	900	1200			1100	1300

the CDC–pyrocarbon composites prepared from SiC. The dynamics of adsorption of benzene vapor by the CDC–pyrocarbon composites prepared from Mo_2C was investigated in [35]. Based on the results, a capillary-condensation mechanism was proposed for this process. A study of the capillary impregnation dynamics showed that the amount of liquid benzene passed through a unit cross-section of the material is proportional to the square root of time [36].

2.5. Analysis of CDC by X-Ray Methods, Electron Microscopy, and Raman Spectroscopy

In Mohun's work [3, 12, 13], graphitization degree (i.e., average size of graphite crystallites) of the CDC was determined by means of XRD. Significant graphitization was only observed in samples obtained at higher temperatures. The diameter of crystallites in the sample produced from SiC at 1650 °C was about 4 nm. The graphitization degree also depended on the duration of the chlorination process: samples produced at 1050 °C (24 hr) and 825 °C (630 hr) demonstrated the same graphitization degree. CDC samples produced at lower temperatures were completely amorphous and showed absolutely no diffraction pattern.

The structure of CDC was also investigated in the early work [37] by XRD. CDC for that study was produced in the course of neutron activation analysis of silicon carbide [38]. β-SiC was treated in flowing chlorine in a quartz tube reactor at 1200–1250 °C. Some samples were slowly cooled after the chlorination and stored for several months prior to XRD measurement. Debye–Scherrer films of them showed only the diffuse rings typical for amorphous substances. Another sample of CDC prepared for XRD was fresh (2 days old) and was rapidly cooled after chlorination in a stream of cold nitrogen. Debye–Scherrer films showed, in addition to the diffuse rings, a number of spotty lines of which most were identified as β-SiC. The samples were easy to break under the touch of fingertips, while the quenched samples were hard and brittle. It is notable that the size and shape of the bodies did not change during the reaction.

L_a and L_c values for heat-treated CDC from TaC were calculated from the corresponding XRD patterns in [14].

Table 7. Properties of CDC–pyrocarbon composites.

Apparent density (g/cm^3)	0.7–0.87
Total pore volume (% vol.)	55–80
Volume of nanopores (% vol.)	35–50
Specific surface area (m^2/g)	700–1390
Specific capacitance (F/cm^3)	30–39
Compressive strength (kg/cm^2)	92–102
Carbon content (% mass)	99.1–99.6

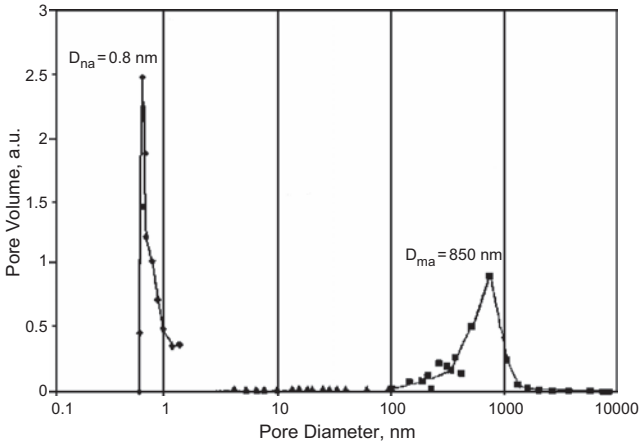


Figure 6. Pore size distribution in the nanoporous article prepared on the basis of TiC derived carbon [32, 137]. ■—mercury porosimetry data; ▲—gas sorption data, BJH method; ♦—gas sorption data, Horvath–Kawazoe method.

The L_a dimensions were larger than for conventional activated carbons treated to the same temperatures. In contrast, the L_c dimensions were very low, as presented in Figure 7. Distinct graphite reflexes became apparent after heat treatment of the CDC from SiC at 2000 °C.

According to the authors' opinion [14], as-prepared CDC from SiC and TaC consists of rather extended layers with very little parallel stacking and the layers must be warped and bent. Upon heat treatment, the first effect is a pronounced increase in the diameter of the coherently scattering segments of the layers. Consequentially, the pore volume accessible to the adsorptives increases. The open structure begins to collapse at 1500 °C as the stacks of carbon layers grow in height and the distance between these stacks shrinks. The pore volume tends toward zero at 1900 °C, and at 2800 °C the layers align to form the graphite structure.

In 1997 Boehm claimed the first observation of carbon nanotubes from electron micrographs taken in 1972 of a

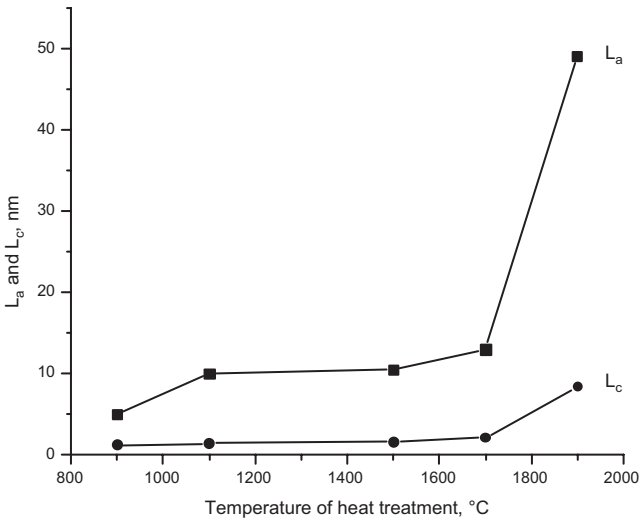


Figure 7. Dimensions of graphite crystallites in the heat-treated CDC produced from TaC at 500 °C [14].

sample of CDC prepared by chlorination of silicon carbide at 800–850 °C [39]. The image of the nanotubes was published in 1978 in the paper by Millward and Jefferson [40]. However, such features could be observed only in a few particles and were considered atypical for this carbon material. Boehm assumed that the carbon nanotubes were not produced during the chlorination reaction, but might be present in the silicon carbide precursor. However, our experience with SiC-derived carbon shows that the formation of multiwall nanotubes (MWNTs) is fairly common and possibly catalyzed by impurities [1].

Further investigations of CDC structure were reported in a number of papers by Gordeev et al. [41–46]. A small-angle X-ray scattering (SAXS) study of the structure of CDC prepared from 6*H* SiC single crystals, polycrystalline α-SiC [41], and polycrystalline SiC, TiC, and Mo_2C [42] produced by treatment in chlorine gas at 700–1000 °C was conducted. The measurements were carried out both in reflection and transmission. The obtained angular dependences of the scattering intensity obtained were treated as a result of scattering from nanoparticles of different size. By decomposing the experimental curves into components corresponding to particles with different gyration radii R_g, scattering distribution functions with respect to gyration radius were found. It was shown that particles with $R_g \approx 1.0$–1.2 nm were the largest fraction in CDC from α-SiC and a 6*H* SiC single crystal [41]. A considerable fraction of 3–6 nm particles was also found in CDC from the 6*H* SiC single crystal. A possible cause of the larger particles' formation in the case of single crystals is their lower chlorination rate and consequently longer treatment time. According to [42], particles with $R_g \approx 0.5$ nm made up the largest fraction in porous carbon, irrespective of the type of the starting carbide. However, samples prepared from different carbides differed in the degree of nanoparticle uniformity. The most uniform in size were nanoparticles in the samples prepared from SiC, in which, on average, R_g was less than 0.6 nm. Nanoparticles in the porous carbon produced from Mo_2C were about twice as large.

The SAXS studies of nanoporous carbon samples obtained from polycrystalline α-SiC, TiC, and Mo_2C and a 6*H* SiC single crystal suggested that the nature of the CDC structure was fractal [43]. Two systems of nanoclusters observed in the CDC structure differed in the type of stacking of their structural elements: small-scale mass fractals of a dimension $1 < D_2 < 3$ and a size $L_2 = 5$–9 nm, which depended on the type of the initial carbide, and large-scale nanoclusters having a size $L_1 > 55$ nm. In most samples, large-scale nanoclusters can be regarded as objects with a fractal surface and a dimension $2 < D_1 < 13$, which also depends on the type of the initial carbide. Large-scale nanoclusters in CDC obtained from Mo_2C proved to be mass fractals with a dimension $D_1 > 2$. The peculiarities of the structural formation of nanoporous carbon obtained from various carbides were also discussed.

The XRD and SAXS data on the CDC structure from [41–43] were revised and summarized with confirmation from high resolution transmission electron microscopy (HRTEM) images in the report [47]. It is notable that, in addition to amorphous and partially graphitized carbon,

some diamond nanocrystallites were detected in the CDC from SiC.

XRD, SAXS, and X-ray photoelectron spectroscopy were used to study the structure of CDC powders produced by chlorination of polycrystalline α-SiC powders with various degrees of dispersion (8–10, 180–200, and 800–1000 μm) [44]. SAXS data were used to derive the distribution function for scattering inhomogeneities with respect to gyration radii, $m(R_g)$. It was shown that the mean size and the fraction of large particles both increase with increasing grain size in the starting powder. As shown by the diffraction patterns, the degree of graphitization of the nanocluster structure increases simultaneously, which is attributed to the longer times required to carbonize coarser grains. An analysis of photoelectron spectra for 1*s* electrons of carbon atoms showed that, for most of the C–C bonds (>65%), the hybridization of valence bonds is intermediate between those for graphite and diamond (sp^x, where $2 < x < 3$), which is indicative of the bending of graphene-like layers in nanoclusters.

A study of the temperature dependences of the electrical resistivity, Hall coefficient, and thermoelectromotive force in the temperature range 1.5–300 K for the materials characterized by SAXS in [43] showed that the structural units responsible for the character of charge transport in these materials were carbon nanoclusters of 1–3 nm in size [45]. The conduction in all the samples studied was found to be of hole type with a high concentration ($n_h \sim 10^{20}$ cm^{-3}) of current carriers.

A new mechanism of nanopore formation in carbon materials produced by the interaction of carbides with chlorine was proposed in [46]. In essence, this method is the following. A series of nonlinear chemical reactions proceed during a chemical interaction between chlorine and carbide. If the external parameters, component fluxes, and diffusion rates satisfy certain relations, a self-organization process can occur, which results in the creation of a periodic nanoporous structure in the carbon material formed. A mathematical model was proposed, the main characteristics of the process were calculated, and the restrictions on the parameters at which the formation of the porous structure becomes possible were found.

The chlorination process of titanium carbide and the obtained CDC were extensively investigated with the use of HRTEM and XRD [20, 48]. HRTEM images of carbon fragments from TiC treated in chlorine at 900 °C revealed that the major part of the carbon structure consisted of three-dimensionally disordered graphitic planes [48]. However, "onionlike" structures were often seen at particle edges. These structures were represented by concentric spherical graphitic shells, interpenetrated and overlapping each other. The outer diameter of the shells was in the range 15–35 nm. The distance between graphitic planes in the shells was 0.34 nm, larger than the interplanar distance in graphite. The carbon sample from TiC had a very high specific surface area (1400 m^2/g) and pore volume and a statistically narrow pore size distribution. As a comparison, HRTEM images of CDC produced at 1000 °C from SiC did not indicate onion-like shells. However, our work showed the formation of carbon onions in CDC produced from SiC [5]. The mechanism of their formation via diamond nanoparticles is described in [49]. HRTEM images of carbon onions and other forms of carbon found in CDC are presented in Figure 8.

It was found that the presence of transition-metal chlorides in the reaction mixture significantly influenced not only the porous structure of CDC from TiC but its degree of graphitization as well [20]. The degree of graphitization was calculated from the XRD pattern according to the empirical formula

$$q = \frac{I_{002}/I_{100}}{14.3} \cdot 100\% \qquad (7)$$

where I_{002} and I_{100} are the relative intensities of graphite (002) and (100) reflections in the diffraction pattern; 14.3 is the I_{002}/I_{100} ratio in high density pyrolitic carbon without an amorphous constituent. As presented in Table 2, the substantial difference between the two series of samples was observed at chlorination temperatures above 700 °C. On the basis of these data, confirmed by HRTEM images, the authors drew the conclusion that the chlorination of TiC generally yields an amorphous carbon below 1000 °C, while at higher temperatures clusters of graphitized carbon appear in the amorphous carbon matrix. The addition of small amounts of the catalytically active metal chlorides of the iron subgroup leads to the formation of better organized carbon structures such as turbostratic and graphitic lamellae at temperatures as low as 800–900 °C.

The structure of CDC produced from aluminum carbide at 300–900 °C was examined in [21, 22]. Degree of graphitization (q) was calculated from XRD patterns according to the method used in [20] and is presented in Table 3. XRD data were confirmed by HRTEM. The CDC sample prepared at 300 °C consisted mainly of amorphous carbon, while increased amounts of ordered structures were observed in materials produced at 500 °C. The CDC from Al_4C_3 produced at 700 °C consisted, to a large part (~80%), of barrel-like multiwall nanoparticles and nanotubes. The sample chlorinated at 900 °C was primarily turbostratic carbon forming randomly oriented bandlike graphitic lamellae and large sheets. The observed nanobarrels were considered to be intermediates between empty onionlike particles and multiwall nanotubes and were patented as a "nanotubular carbon material" [50].

Raman spectra of CDC prepared from polycrystalline α-SiC and TiC were reported in [51]. The spectra were shown to be of a multicomponent nature, which distinguishes this group of materials from graphites and disordered carbon structures. A series of low-frequency modes were detected and anisotropy and dispersion effects were revealed. The results obtained argue in favor of a cluster structure of nanoporous carbon and size quantization of the electronic and vibrational spectra in carbon nanoclusters. The main structural elements of the nanoclusters in the porous framework were small fragments of bent or broken graphene sheets. The presence of fragments of strained cubic or hexagonal diamond was observed in a number of cases.

Formation of carbon films and carbon-rich surface layers was repeatedly observed during the investigation of high-temperature corrosion of SiC-based ceramics in mixed oxygen–chlorine environments [52–58] and the etching of

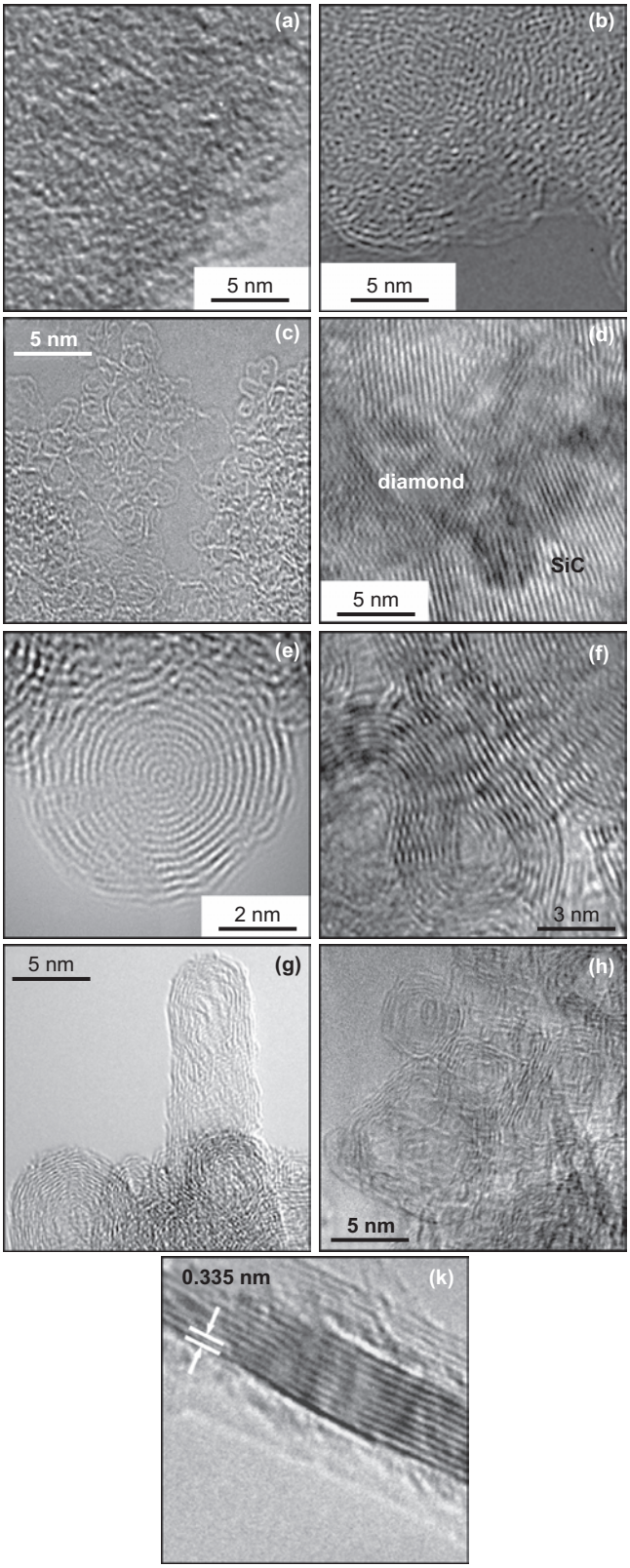


Figure 8. HRTEM images of various forms of carbon found in CDC samples. (a) Amorpous carbon; (b) disordered graphite; (c) fullerene-like carbon; (d) nanodiamond; (e) onion; (f) ribbons; (g) MWNT; (h) barrel-like particles; (k) planar ordered graphite.

SiC surfaces with chlorine and fluorine-containing gases for device fabrication [59–62]. The properties and structure of the formed carbon were not the object of cited studies; moreover, its formation was frequently considered to be undesirable.

Carbon coatings on silicon carbide powders and monolithic specimens produced by treatment in diluted chlorine and chlorine–hydrogen mixtures were examined by Gogotsi and McNallan in a series of works [1, 2, 5, 10, 63–65]. β-SiC powder was treated in Cl_2(2–3.5%)–H_2(0–2%)–balance Ar gas mixtures in the temperature range 500–1000 °C at ambient pressure [10, 63, 64]. Experiments were continued for 5–72 hours depending on the reactivity. The progress of the reactions conducted at 800 °C was monitored by thermogravimetric analysis. The mass decreased approximately linearly with time until it approached a minimum value which was dependent on the ratio of chlorine to hydrogen. Gas mixtures that were low in hydrogen produced the highest reaction rates. Uniform amorphous carbon or nanocrystalline graphite films, as shown by XRD, Raman spectroscopy, and TEM, were produced by treatment in Cl_2–Ar media. In the Raman spectra, D- and G-bands of disordered graphite were observed. From the relative intensities of these bands the size (L_a) of the graphite crystallites was evaluated to be 2–4 nm as presented in Table 8. As the treatment temperature increases, the size of these crystallites and the thickness of the film formed after a fixed period of exposure both increase. BET specific surface areas of carbon layers were calculated using the results of N_2 adsorption measurements and taking into account the presence of unreacted silicon carbide (Table 8). High values of surface area supply evidence that nanoporous carbon coatings were produced.

The addition of hydrogen into the reaction mixture not only reduced the reaction rate but influenced the composition and properties of the products. The sharp diamond band at ca. 1330 cm^{-1} was observed in the Raman spectra of the powder treated in 2.6% Cl_2–1.3% H_2–Ar at 950 °C, and diamond regions were clearly visible under an optical microscope. Other bands which may correspond to crystallized and amorphous diamond were present in the Raman spectra as well. Thus, hydrogen plays a significant role in the formation of diamond by stabilizing the dangling bonds of the cubic diamond formed from the structure of β-SiC upon the extraction of silicon.

Carbon coatings prepared by chlorination of other carbide materials including TiC [11, 66] and TaC [11] were so far reported to consist of disordered graphitic or amorphous carbon without the presence of sp^3 hybridized species.

Chlorination of SiC fibers was also studied [67, 68]. Fibers from polymer precursors are nanoscale mixtures of SiC,

Table 8. Specific surface area and size of crystallites of CDC powders.

Media	T (°C)	t (hrs)	S_{BET} (m^2/g)	L_a (nm)
3.5% Cl_2–Ar	600	27	400	2.0–2.5
	600	48	1163	2.4–2.6
	1000	20	1010	2.0–4.5
2% Cl_2–2% H_2–Ar	1000	72	10	1.0–4.5

carbon, and SiC_xO_y. Their chlorination can be conducted at 400 °C and above, well below the temperatures necessary to produce CDC on SiC crystals or micropowders.

Treatment of monolithic sintered and CVD samples of α- and β-SiC under the conditions described previously [10] resulted in the formation of gray, translucent (in thin sections) films in the cases of low Cl_2:H_2 ratios in the reaction mixture [1, 2, 5, 65]. These nonporous films probably acted as a diffusion barrier and prevented further reaction. Lattice fringing, electron diffraction patterns, and electron energy-loss spectroscopy confirmed the formation of nanocrystalline diamond-structured carbon. The analysis of Raman spectra supported the presence of sp^3 bonding, which is also confirmed by the very high hardness (up to 50 GPa) and Young's modulus (up to 600 GPa) of these coatings.

CDC coatings produced by treatment in chlorine or chlorine–hydrogen mixtures at Cl_2:H_2 ratios equal to or larger than 2:1 consisted of a loosely attached, powdery top carbon layer with a dense, highly disordered, adherent amorphous or nanocrystalline layer below it. The adherent layers were shown to be composed of a mixture of graphitic and diamond-structured carbon [5] or highly disordered sp^2-carbon [2]. The thickness of the carbon layer increased linearly during the process. Because the kinetics is linear, the controlling factor of the reaction is not the diffusion of reactant species through the growing carbon layer. Therefore, the layer is nanoporous.

A theoretical study of graphite and diamond growth on the (1000) Si surface of 6*H* SiC was carried out using molecular dynamics simulation [1, 69]. It was shown that very high lattice strains do not allow the continuous growth of diamond on SiC, and fragmentation occurs that leads to nanocrystalline material. Small diamond clusters on SiC are predicted to show good adhesion to the substrate and maintain sp^3 coordination of carbon atoms in the cluster. A TEM study [5] proved that SiC was converted to sp^3 carbon, and formation of crystalline diamond-structured carbon occurred from disordered Si-depleted SiC within nanometers of the SiC/C interface. Growth of larger crystals of diamond-structured carbon was probably the result of the coalescence of continuous nanocrystalline regions. The presence of hydrogen helps to maintain the sp^3 hybridization of carbon by stabilizing the dangling bonds.

A detailed HRTEM study of diamond crystals produced by treatment of SiC in the gas mixtures (1–3.5%)Cl_2–(0–2%)H_2–Ar at 1000 °C showed that growth of diamond under these conditions occurred epitaxially on the SiC surface [49] (see Fig. 8d). It has been shown that the presence of hydrogen in the environment is not required for diamond synthesis. However, hydrogen can stabilize the nanocrystals and lead to the formation of thick diamond films. If no hydrogen is added, diamond nanocrystals transform to graphitic carbon, forming carbon onions and other curved graphitic structures. Diamond formation is also favored by the low synthesis temperature providing moderate carbon mobility. This result is in accordance with X-ray photoelectron spectra of CDC from SiC, TiC, and ZrC produced at various temperatures [70], in which it was shown that the degree of sp^3 hybridization increases as the process temperature is reduced from 1100 to 350 °C.

3. SELECTIVE ETCHING OF CARBIDES BY MELTS AND SUPERCRITICAL WATER

3.1. Reaction of Calcium Carbide with Inorganic Salts

Saltlike carbides react at elevated temperatures with a number of reagents resulting in carbon formation along with other reaction products [71]. Infrequent reports on this carbide-thermal process and the resulting carbon material were all devoted to calcium carbide reactions with the metal salts [72–78]. Generally, the process can be considered as a high-temperature exchange reaction of calcium carbide and metal salt, producing a calcium salt and a new unstable saltlike carbide that decomposes to give free metal and carbon [77].

This method was first applied for the production of macrocrystalline graphite with the use of magnesium chloride [72]. The process included the following exchange and decomposition reactions:

$$CaC_2 + MgCl_2 = MgC_2 + CaCl_2 \quad (8)$$

$$2MgC_2 = Mg_2C_3 + C \quad (9)$$

$$Mg_2C_3 = 2Mg + 3C \quad (10)$$

The reaction mixture was heated to 600–1275 °C for 1–5 hours. The best yield and quality of graphite were achieved in the range 950–1200 °C. After the removal of the unreacted carbides, salts, and amorphous carbon the yield was as much as 85%. XRD showed that both graphitic and turbostratic carbon were present in the product. It is important to note that since $CaCl_2$ has a high melting point (772 °C) and low vapor pressure, pure CDC cannot be produced by chlorination of CaC_2 (Fig. 2b).

The porous structure and adsorption properties of the CDC material prepared by the carbide-thermal process were first studied by Fedorov et al. [73]. The reaction of calcium carbide with sodium chloride according to the overall equation

$$CaC_2 + 2NaCl = 2C + 2Na + CaCl_2 \quad (11)$$

was conducted in the temperature range 400–900 °C and resulted in a sorption-active carbon material. The parameters of the porous structure were determined by measuring benzene sorption isotherms. As one may see from Table 9,

Table 9. Parameters of the porous structure of CDC from CaC_2 + NaCl.

Temperature (°C)	V_{na} (cm^3/g)	V_{me} (cm^3/g)	S_{me} (m^2/g)
400	0.13	0.50	332
500	0.22	0.24	102
600	0.11	0.40	198
700	0.00	0.63	253
900	0.04	0.12	55

Note: V_{na} and V_{me} are volumes of nano- and mesopores, respectively; S_{me} is the specific surface area of mesopores.

carbonaceous materials from CaC_2 have extensive and variable mesopore volume but a lower specific surface area compared to CDC produced by chlorination of carbides.

Measurements of the free carbon yield at various CaC_2:NaCl ratios in the initial mixture showed that two reaction pathways are possible at 900 °C [76]. The chloride-catalyzed decomposition of CaC_2 takes place at low CaC_2:NaCl ratios (1:0.03–1.2 eq.). Reaction (11) proceeds at higher CaC_2:NaCl ratios.

The technique for producing CDC adsorbents from CaC_2 includes the following steps: mixing of the initial components with the particle size of 100–300 μm in the ratio 1:(1–3), compaction at 5–40 MPa, thermal treatment at 600–1000 °C followed, if necessary, by a water-acid treatment, drying, and crushing. The properties of the CDC adsorbents prepared by reaction of CaC_2 with a variety of chlorides are summarized in Table 10. These sorbents can be used as soil detoxifiers, gas drying agents, chemical adsorbents, and catalysts [78].

3.2. Hydrothermal Leaching of Carbides

3.2.1. Thermodynamic Analysis

The thermodynamic stability of a number of metal carbides under hydrothermal conditions was analyzed in [79]. The possible corrosion routes and products were predicted using Gibbs energy minimization sofware. Thermodynamic calculations were conducted for the temperature range from 300 to 1000 °C and pressures of 2 and 200 MPa for the M_xC–nH_2O(g) system, where $n = 2$ and 10 and $x =$ Si, Ti, Ta, Nb, W, and B. The following idealization was supposed from these calculations: hydrothermal fluid is treated as a steam, the solid phase is a mechanical mixture of compounds with unit activities, and the gas phase is a mixture of ideal gases. Furthermore, kinetic factors and effects such as pH were not taken into consideration.

The analysis indicated that carbon, MeO_x, CH_4, CO_2, CO, and H_2 are the main products of hydrothermal corrosion in the temperature range of the study. The principal reactions in the M_xC–nH_2O(g) system can be written as follows:

$$x/2\mathrm{MeC} + x\mathrm{H_2O} = \mathrm{Me}_{x/2}\mathrm{O}_x + x/2\mathrm{CH_4} \quad (12)$$

$$\mathrm{MeC} + (x+1)\mathrm{H_2O} = \mathrm{MeO}_x + \mathrm{CO} + (x+1)\mathrm{H_2} \quad (13)$$

$$\mathrm{MeC} + (x+2)\mathrm{H_2O} = \mathrm{MeO}_x + \mathrm{CO_2} + (x+2)\mathrm{H_2} \quad (14)$$

Table 10. Parameters of the porous structure of CDC from CaC_2.

Chloride	V_{na} (cm^3/g)	V_{me} (cm^3/g)	S_{me} (m^2/g)	W_0 (cm^3/g)	E_0 (kJ/mol)
LiCl	0.08	0.35	95		
KCl	0.26	0.71	270	0.25	19.7
$MgCl_2$	0.11	0.41	115		
$AlCl_3$	0.02	0.46	127		
$SnCl_2$	0.27	0.42	94		
$PbCl_2$	0.39	0.49	111		
$ZnCl_2$	0.21	0.58	382	0.20	18.6
$MnCl_2$	0.14	0.41	101	0.15	14.8
$CoCl_2$	0.16	0.38	95	0.17	18.6
$NiCl_2$	0.14	0.24	87	0.13	19.2

$$\mathrm{MeC} + x\mathrm{H_2O} = \mathrm{MeO}_x + \mathrm{C} + x\mathrm{H_2} \quad (15)$$

$$\mathrm{MeO}_x + n\mathrm{H_2O} = \mathrm{MeO}_x \cdot n\mathrm{H_2O} \quad (16)$$

These calculations show that the formation of carbon on the surface of SiC depends crucially on the H_2O:SiC ratio. Three regimes can be identified:

1. At low H_2O:SiC ratios, both carbon and silica are deposited on the surface.
2. At intermediate H_2O:SiC ratios, carbon and silica are formed. Silica is dissolved in water according to eq. (16). Thus, reaction (15) tends to move further to the right.
3. At higher H_2O:SiC ratios (ca. 10) neither carbon nor silica precipitates, dissolving in the excess of water.

It is shown that the yield of methane, which is unstable at high temperatures, decreases with temperature and thus the yield of elemental carbon produced due to the instability of CH_4 increases. Higher pressures lead to a slight increase in methane stability and, accordingly, to a lower yield of carbon.

TaC, TiC, and NbC show behavior which is very similar to that of SiC. In particular, TaC seems to be more stable at low amounts of water. Carbon formation is calculated for TaC at temperatures above 700 °C and only at relatively low pressures. It is important to remember that most metal oxides, unlike silica, cannot be easily dissolved in supercritical water. Thus, a mixture of carbon and metal oxide is produced. The calculated behavior of WC is somewhat different as compared with previously discussed carbides. The carbide W_2C and a number of stable oxides may form. Free carbon is predicted to form only at the lower temperatures. Boron carbide is the only carbide from the group studied that is not predicted to convert to free carbon under hydrothermal conditions. These calculations provide some general guidelines and, as will be shown, are generally in good agreement with experimental results.

The formation of carbon-rich layers on the surface of carbides was repeatedly detected in the course of investigations on the controlled oxidation of HfC [80–85], ZrC [84–89], and TiC [83, 85] at low oxygen pressures (0.02–16 kPa). These layers contained up to 25 at% carbon in addition to metal oxides and oxicarbides. This carbon was mainly amorphous as characterized by Raman spectroscopy. In some cases, the formation of carbon of the hexagonal diamond structure was observed [86].

3.2.2. Experimental Results

The hydrothermal formation of carbon from carbides was first reported in 1994 by Gogotsi and Yoshimura [64, 90–93]. This phenomenon was observed during corrosion studies performed on amorphous SiC (Tyranno) fibers at 300–800 °C and about 100 MPa. Tyranno fiber is a type of silicon carbide fiber which contains titanium and oxygen and is made by pyrolysis of an organometallic polymer precursor. It consists of SiC nanocrystals which are connected by Si–O–C and Ti–O (Ti–C–O) interlayers; the structure can be considered as an amorphous SiO_xC_y matrix with nanometer-size β-SiC inclusions.

The outer surface of the fibers treated at temperatures of up to 450 °C remained smooth and almost featureless. Auger electron spectroscopy depth analysis of the surface layer showed that a carbon layer containing traces of silicon was formed [91]. The thickness of this layer varied from 10–20 nm to 1–2 μm. Raman spectra of the fibers showed the graphite (G) and disorder-induced (D) bands, thus confirming the formation of amorphous or nanocrystalline carbon [90]. The crystallites were of very small size and sp^2 C–C bonds dominated.

The formation of amorphous carbon on the surface of Tyranno fibers after hydrothermal treatment at 300–400 °C and 100 MPa was also reported in [64, 67, 94, 95]; on the surface of CVD-SiC fibers after treatment at 400–700 °C and 200 MPa [96]; and on the surface of α- and β-SiC platelets, whiskers, and powder at 600–800 °C [64, 79, 94, 95, 97]. Some amount of disordered diamond, as well as graphite and amorphous carbon, was found on the surface of α- and β-SiC powder and single crystals after treatment at 300–800 °C and 100–500 MPa [98–102]. The hydrothermal method was also found to be suitable for the preparation of CDC from other carbides including WC, TaC, NbC, and TiC at 500–750 °C and 100–170 MPa [79].

3.3. Carbides as a Source of Carbon in High Pressure–High Temperature Synthesis of Diamond

Diamond has been synthesized from graphite at high pressure and high temperature (HPHT) using some transition metals as a solvent catalyst. This process is accompanied by the formation of stoichometric and nonstoichometric metal carbides. The role of metal carbides in the diamond formation process was disputed [103, 104]. According to the suggested mechanism, a stoichometric iron or chromium carbide with higher carbon content decomposes into the carbide with a lower carbon content to precipitate diamond.

A HPHT process of diamond formation from carbides in the presence of metal oxides or halogenides was patented in 1988 [105, 106]. Diamond powders can be prepared from carbides of silicon, calcium, aluminum and other elements at pressures >1 GPa in the temperature range 400–1200 °C according to the general reaction equation

$$a\mathrm{MeC}_x + \mathrm{Me'X}_{y/a} = a\mathrm{MeX}_y + \mathrm{Me'C}_z + (ax - z)\mathrm{C}_{\mathrm{diam}}$$

where X = O, F, Cl, Br; $0 \le z < ax$. For example, diamond was produced after treatment of a stoichometric mixture of Al_4C_3 and MoO_2Cl_2 at 1.5 GPa and 400 °C.

The first detailed description of the HPHT synthesis of diamond from silicon carbide in the presence of metal solvent-catalyst was published by Hong et al. [107, 108]. The SiC–Co systems with a weight ratio of 1:8 were treated at pressures of 4.5 or 5.5 GPa and temperatures of 1400–1550 °C. At a pressure of 5.5 GPa and temperatures of 1430–1550 °C, the SiC decomposed and diamond and graphite appeared. Using the Ni–Mn–Co alloy as a solvent catalyst was proved to be more effective in the synthesis of diamond from SiC [109]. Diamond formed rapidly in the SiC–$Ni_{70}Mn_{25}Co_5$ system at a pressure of 5.4–6.0 GPa and temperature of 1350–1570 °C. The carbide was completely decomposed during 6 minutes of HPHT treatment. After oxidizing the graphite in a hot mixture of H_2SO_4 and HNO_3, the yield of diamond powder was 56%. It was shown that the mechanism of diamond formation in the SiC–metal system is different from that in the graphite–metal one. First, SiC decomposes into carbon and silicon, and then carbon, by the action of the metal catalyst, is separated into diamond and graphite while silicon may dissolve in the metal. HPHT synthesis of diamond from boron carbide was conducted in the work [110]. A significant amount of diamond grains was obtained at pressures ≥5.8 GPa and temperatures ≥1460 °C.

The formation of diamond films and particles was observed in the course of iron carbide decomposition under HPHT conditions. Diamond films were heteroepitaxially grown on the surface of the cubic boron nitride particles at 5.2 GPa and 1370 °C [111]. Homoepitaxial growth of the diamond films was performed by treating the mixture of Fe_3C and diamond seeds at 5.6 GPa and 1430–1470 °C [112]. Decomposition of high-purity Fe_3C powder at 5.3 GPa and 1360 °C resulted in diamond particles. These experiments have also confirmed the possibility of intermediate carbide formation in the HPHT process of diamond formation in the graphite–metal systems.

4. THERMAL DECOMPOSITION OF CARBIDES

It is known that silicon carbide, as well as many other carbides, decomposes or melts incongruently at high temperatures [71]. This thermal decomposition in vacuum or an inert environment is characterized by the formation of carbon and the melting and evaporation of metal, because the melting point of graphite exceeds that of most metals.

Carbon formed after the heat treatment of SiC in vacuum at temperatures up to 2000 °C approximately was first mentioned in 1960 in the discussion attached to [37]. Just as in the case of carbon material prepared from silicon carbide by selective removal of silicon by chlorine gas, a pseudomorph of the original carbide sample was observed. In other words, the size and shape of the original particles did not change during the reaction. The same phenomenon was observed in 1963 during of the thermodynamic study of SiC dissociation [113].

A few attempts to investigate the structure of carbon obtained by the decomposition of silicon carbide were reported in the following decades. An XRD study was conducted by Badami [114, 115] in which single crystals of α-SiC ($6H$) were heated to temperatures of about 2050–2150 °C for 1 hour at a pressure of 10^{-5} Torr [114]. The sample produced by heating to 2050 °C had undergone only partial decomposition as demonstrated by its diffraction pattern. On the basis of an examination of various parts of the sample produced at 2150 °C, the following structure was proposed. The pseudomorph contained two distinct structures: a matrix of turbostratic carbon with a strong preferred orientation of the carbon layers parallel and perpendicular to the layers of silicon carbide, and within this matrix there were crystalline regions of graphite having a common c axis and axial symmetry about it. X-ray data on the pseudomorph from single crystals of $6H$ silicon carbide decomposed at

2280 °C in vacuum for 1 hour showed that an outer layer of turbostratic graphite formed, and deeper inside the sample were single crystals of graphite whose orientation was also determined by the structure of SiC [115]. Further into the interior, turbostratic graphite was found again, and, finally, completely disordered carbon. A possible mechanism for the decomposition was the collapsing of two layers of carbide lattice to give one layer with a hexagonal arrangement of carbon atoms. Molecular dynamics modeling shows that graphite layers can be accommodated on the surface of SiC [1]. Electron diffraction studies detected the formation of graphite on the surface of a SiC crystal at temperatures as low as 1200 °C.

Selected area electron diffraction was used to investigate the decomposition of single-crystal β-SiC whiskers [116]. A cluster of the whiskers, averaging about 300 nm in diameter, was heated to about 2400 °C at 10^{-6} Torr for 2–3 min. The whiskers were found to be in various stages of decomposition. Graphite crystallites about 10 nm or less in diameter were observed at the outer surface of the partially decomposed whiskers. Based on the selected area diffraction patterns, it was concluded that the crystallites grew with their basal planes tangent to the whisker surface. No evidence was found to indicate that the formed carbon was structurally related to the parent silicon carbide.

Thermal decomposition in vacuum (10^{-10} Torr) of Si- and C-faces of 6*H* SiC(0001) crystals was studied by low energy electron diffraction and Auger electron spectroscopy [117]. The formation of a graphite monolayer started at 800 °C. On the Si-face this monolayer was monocrystalline with the unit mesh rotated 30° with respect to the SiC unit cell. On the C-face this layer was polycrystalline with the unit mesh randomly oriented. The authors proposed that the collapse of the carbon of three successive SiC layers is the most probable mechanism for the initial stages of the graphitization of SiC basal planes. These results agree well with ellipsometric data on the surface decomposition of 6*H* SiC(0001) crystals [118]. It was also shown that rate of the carbon layer growth on the carbon face is about five time higher than that on the silicon face. Our modeling work showed that a graphite layer can form on the (0001) surface of 6*H* SiC, and good adhesion is expected for the first two monolayers [1].

Iijima studied the decomposition of silicon carbide under electron beam irradiation in 1982 [119]. Small flakes of 6*H* SiC were irradiated with the electron beam at a pressure of 10^{-6} Torr. The decomposition was controlled to some extent by halting the irradiation after a short exposure to the beam. Electron diffraction patterns showed a gradual disappearance of the SiC reflections with simultaneous graphitization of the sample. The preferred orientation of graphite was similar to the X-ray diffraction observations of Badami [114]. However, the decomposition process was not completely topotaxial, which is confirmed by electron micrographs. Many ordered graphite bands and no regions of amorphous or turbostratic carbon were recognized in these micrographs. Images of the interface between SiC and graphite clearly showed the topotaxial mechanism of the reaction.

A consideration of transport processes leading to the formation of the carbon pseudomorph during the thermal decomposition of SiC was attempted by Motzfeldt and Steinmo [120]. The vapor pressure over the SiC(s)–C(s) was studied by means of the Knudsen effusion method for three different orifice sizes. It was found that surface diffusion is of great importance in the process. The overall process was pictured in the following way: the outermost layer of the SiC crystal loses silicon by evaporation from the surface, and a layer of microporous carbon is left. A gradient in silicon is established through this layer and silicon from the interior moves through it by surface diffusion and then evaporates from the surface. One can find indirect evidence on silicon surface migration during the silicon carbide decomposition in the report from 1970 [121].

The only published source on the thermal decomposition of carbides other than SiC is [122]. Decomposition of aluminum carbide was carried out by heat treatment at 2400 °C and atmospheric pressure for 2 hours in argon. Formed graphite crystals had a low *c*-spacing (0.33545 nm), chemical purity, large size, and general absence of cleavage and twinning.

4.1. Synthesis of Carbon Nanotubes

A new interest in the field of the thermal decomposition of silicon carbide arose in 1997 when Kusunoki et al. discovered growth of self-organized carbon nanotube (CNT) films during the process [4]. The β-SiC (3*C*) particles were heated to about 1700 °C at 2×10^{-5} Pa for a few minutes in a TEM equipped with a YAG laser heating holder. Overlapped carbon nanotubes with caps of 2–5 nm oriented along the [111] direction on the (111) surface plane of β-SiC single crystal were found (Fig. 9) By heating the specimen to 1900 °C, the shape of CNTs became sharp and they began to grow in various directions. Although the authors did not claim SWNT formation, the image in Figure 9 suggests that single-wall nanotubes were present.

The possibility of the transforming of SiC to nanotubes on larger SiC particles by means of heating the particles in

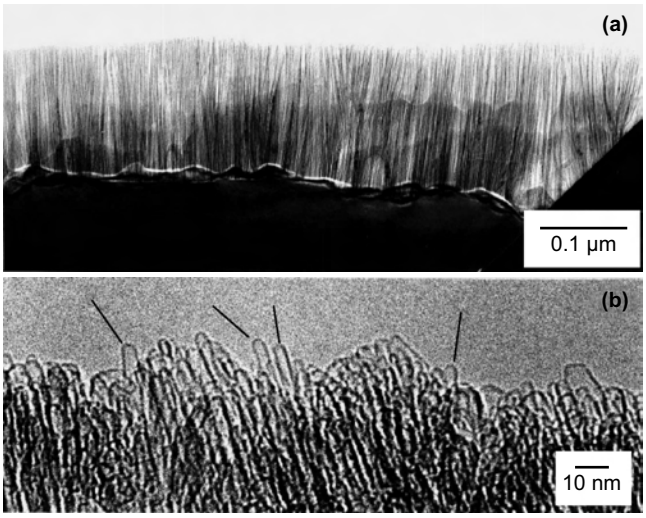


Figure 9. (a) TEM image of the surface of a decomposed α-SiC wafer. (b) Top of the nanotube layer in (a). Arrows show tube caps typical of a SWNT. Reprinted with permission from [124], M. Kusunoki et al., *Jpn. J. Appl. Phys.* 37, L605 (1998). © 1998, Publication Board, Japanese Journal of Applied Physics.

a vacuum resistance furnace was investigated in [123]. SiC powder was heated to 1600–1700 °C for 10 to 15 minutes on a tungsten boat. Dense layers of MNWTs without nanoparticles were observed on the surface of 40% of the SiC powder. The typical nanotubes size was 20–50 nm in diameter and up to 1 μm in length. As in the previous work [4] most of the nanotubes tended to grow almost perpendicular to the original surface of the particle.

A large-area nanotube film was prepared on the surface of an α-SiC wafer by heating at 1700 °C and 10^{-4} Torr for 30 min [124]. The nanotubes were 2–5 nm in diameter and 0.15 μm in length, similar to those shown in Figure 9. The aligned nanotube film was formed uniformly all over the surface of a wafer $3 \times 5 \times 0.34$ mm in size. The (0002) lattice distance of the graphite making up the CNTs was found to be 0.344 nm from the electron diffraction pattern.

Kusunoki et al. proceeded with the investigation of surface decomposition of the α-SiC (0001) single-crystal wafers in the temperature range 1300–2000 °C at 10^{-4} Torr for 30 min [125]. When a wafer was heated at 1300 °C, small caps of 2–5 nm in size consisting of a couple of layered graphite sheets were formed all over the SiC crystal. When the wafer was heated at 1500 °C, the length of CNTs increased to 50 nm after 0.5 h and to 180 nm after 6 h of heat treatment. When the wafer was heated at 1500 °C for 0.5 h in a better vacuum (10^{-6} Torr), the length of the CNTs was only 25 nm. This fact indicated that the oxidation process is essential for CNT formation. Finally, when the SiC wafer was heated at 2000 °C for 0.5 h, a graphite layer of 10 μm in thickness was formed. On most of the layer, graphite sheets were oriented parallel to the $(0001)_{SiC}$ plane and, under this condition, CNTs were not observed. HRTEM micrographs oriented along a CNT axis plan-view direction showed that these CNTs are multiwall and contained two to five walls.

An electron energy-loss spectroscopy line-scan profile showed the segregation of oxygen at the interface between the CNT film and SiC plane. This result indicated that residual oxygen in the vacuum chamber plays an important role in the formation of CNTs. Temperature dependence of the CNT size allowed the authors to propose a model of their formation mechanism, which consisted of three stages. First, at temperatures over 1000 °C, several graphite sheets are formed parallel to the $(0001)_{SiC}$ plane by oxidation of SiC. Second, at around 1300 °C carbon nanocaps are formed by the generation of bubbles of a SiO gas, and the graphite sheets start to stand up. Finally, CNTs grow toward the interior of the SiC crystal (Fig. 10).

The initial growth process of CNT from SiC was also investigated by time-resolved HRTEM [126]. A sample of 3*C* SiC single crystal was resistively heated to 1360 °C at a pressure of less than 4×10^{-5} Pa. Dynamic observation was conducted at a time resolution of 1/60th of second. Two types of CNT were observed with diameters of 0.8 and 1.5 nm. It was shown that amorphous carbon was first formed and then crystallized to graphene sheets parallel to the (111) plane. Carbon nanotubes were formed by the lift of a part of the graphene sheet along the [111] direction due to the generation of pentagons and heptagons.

A remarkable difference between the carbon structures on Si(0001) and C(000$\bar{1}$) faces of an α-SiC single crystal was

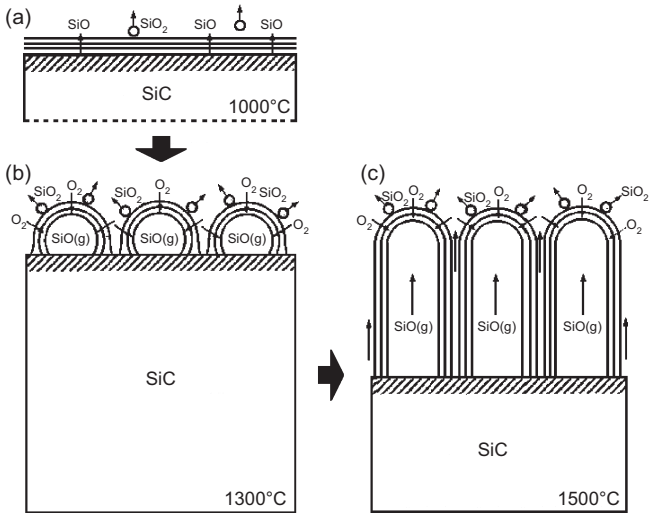


Figure 10. Schematic diagram for the formation mechanism of the CNT film on a SiC crystal by the surface decomposition. Reprinted with permission from [125], M. Kusunoki et al., *Philos. Mag. Lett.* 79, 153 (1999). © 1999, Taylor & Francis.

observed after its decomposition in the temperature range 1200–1700 °C for 30 minutes [127, 128]. Cross-sectional HRTEM of the C-face showed cap formation and CNT growth according to the proposed model, while HRTEM of the Si-face showed that flat graphite sheets were formed parallel to the surface. The difference between C- and Si-faces was confirmed during the process of creating freestanding CNT films by sublimating decomposition of 3*C* SiC(111) films, which were epitaxially grown on Si(111) substrate and afterwards removed [129]. SiC films could be converted to CNT only after removal of the Si- layer.

However, in a later study by Derycke et al., CNTs were found on a Si(0001) face of partially decomposed 6*H* SiC wafers using scanning tunneling microscopy (STM) and atomic force microscopy (AFM) [130]. Heat treatment of the samples at 1650 °C was performed in ultrahigh vacuum ($<10^{-9}$ Torr) by passing a current through them. STM and AFM images clearly showed highly ordered single-walled nanotube (SWNT) networks on and below the surface with their axes parallel to it (Fig. 11). The nanotubes had a very narrow size distribution in the 1.2 to 1.6 nm range.

Nanotubes prepared by surface decomposition of SiC were found to be exclusively of the zigzag type [131]. CNTs were grown on the C(000$\bar{1}$) faces of a 6*H* SiC single crystal wafer by heating it to 1500 °C at a very low heating rate under vacuum (10^{-2} Pa). HRTEM images indicated that the zigzag CNT structure evolved from Si–C hexagonal networks in the SiC crystal by the collapse of carbon layers that remained after the silicon removal. This work is the first report on the selective preparation of carbon nanotubes of the same structure. Previous synthesis methods, such as arc discharge, laser evaporation, and chemical vapor deposition, usually result in zigzag, armchair, and chiral types of CNTs coexisting in the products.

Reduced silicon carbide decomposition temperatures are needed to form carbon nanotubes in the presence of metal catalysts [132]. The metal-coated substrates of polycrystalline

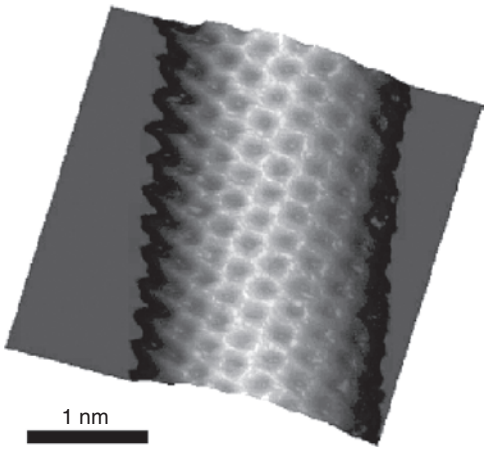


Figure 11. Atomic resolution STM image of a chiral single wall carbon nanotube. Reprinted with permission from [130], V. Derycke et al., *Nano Lett.* 2, 1043 (2002). © 2002, American Chemical Society.

SiC were heated to 1000–1200 °C at 10^{-2} Torr in a vacuum electric furnace. Carbon tubular nanostructures (tubes and fibers) grew in the case of Pt, Ni, Co, Rh, and PtPd at 1200 °C. Scanning electron microscopy observation revealed that the synthesized structures ranged from 100 nm to 10 μm in length. According to TEM images, the synthesized CNTs are MWNTs ranging from 2 to 10 nm in diameter. Most of the tubular structures included a metal particle of corresponding diameter at the tip. The growth mechanism is thought to be similar to that of a CVD process.

Thus, both SWNTs and MWNTs can be formed by the CDC technique with and without catalysts.

5. APPLICATIONS

It has already been mentioned that CDC has a variety of applications including molecular sieves, sorbents for gas chromatography, soil detoxification, dehumidification, and catalysis. This section describes some other potential applications of CDC in more detail.

5.1. Supercapacitors

Carbide-derived carbon is an attractive material for electrodes of electrochemical double-layer capacitors, commonly called "supercapacitors" [133]. Among the advantages of CDC, the high specific surface area and controlled pore size distribution are the most important. Strength and mechanical stability are also necessary for good electrode performance.

Two approaches to the design of the CDC electrode bodies have been proposed. The first one was developed for production of nanoporous carbon articles of any prescribed form and sufficient mechanical strength [30–32] and was employed for the preparation of supercapacitor electrodes [134–136]. The parameters of porous structure and specific capacitance of prepared electrodes measured in aqueous solutions using potassium hydroxide and sulfuric acid as electrolytes are presented in Table 7.

The second approach is more conventional: CDC powder is mixed with an appropriate binder and then formed and treated to achieve mechanical strength [137]. Particular attention in this work was paid to selective control of the resultant pore size in the electrodes to tailor them to fit the ionic size of the positive and negative electrolyte ions employed in the capacitor. Even though the method of manufacturing nanoporous carbon by chlorination of selected carbide or carbide mixtures under appropriate conditions gives good control of the size distribution of nanopores in the resulting carbon particles (see Table 11), some additional refinement of the process control was considered to be desirable. This was achieved by exposing the CDC samples to an oxidizing medium such as water vapor, carbon dioxide, or concentrated nitric or sulfuric acid. The effects of this treatment were controllable widening of the nanopores and removal of physically and chemically adsorbed chlorine. Examples of electrochemical properties before and after modification are given in Table 12.

5.2. Hydrogen Storage

The low atomic weight and density of carbon facilitate great interest in hydrogen uptake and storage in carbon nanomaterials [138]. It is known that hydrogen uptake depends on the porous structure of the adsorbent. The relationship between hydrogen uptake and BET specific surface area for various carbonaceous materials, including activated carbons and nanofibers, is presented in Figure 12 [139]. The experiments on dissociative adsorption of hydrogen on the compact CDC adsorbent prepared according to Gordeev et al. [30–32] were mentioned in [28]. It was shown that the material under study was able to adsorb hydrogen at 300–600 °C and that the adsorbed hydrogen was strongly bonded and could be desorbed at elevated temperatures in the range of 400–1200 °C.

The process of hydrogen sorption by CDC was investigated at temperatures of 300–700 K and pressures of 200–580 atm using the method of thermodesorption spectrometry [140, 141]. Thermodesorption spectra registered in the temperature range of up to 1500 °C consist of several chemisorption-induced phases in several independent states on the surface of the nanoporous carbon. The desorption peaks at 500–550 and 800–850 °C were attributed to the hydrogen adsorbed on dangling carbon bonds and

Table 11. Examples of influence of the carbide precursor on capacitance.

Initial carbide	Nanopore volume (cm^3/g)	S_{BET} (m^2/g)	Specific capacitance (F/g)
Al_4C_3	0.60	1353	251
B_4C	0.78	1782	217
Mo_2C	0.90	1873	223
TiC	0.73	1340	212
SiC	0.44	1059	209
TiC/Al_4C_3, 3/1 (by wt.)	0.58	1542	239
B_4C/Al_4C_3, 3/1 (by wt.)	0.67	1614	239
B_4C/Al_4C_3, 1/1 (by wt.)	0.58	1572	219
B_4C/Al_4C_3, 1/3 (by wt.)	0.53	1440	211

Note: Specific capacitance measured in aqueous 6M KOH at dc = −1 V).

Table 12. Properties of nonmodified and modified CDC materials.

Initial carbide	Type of material	S_{BET} (m^2/g)		Specific capacitance (F/g)	
		powder	electrode	dc = −1.4 V	dc = +1.4 V
SiC	nonmodified	1086	931	5.8	79.9
	modified (H_2O)	2140	1567	92.8	88.1
TiC	nonmodified	1485	1054	80.5	113.4
	modified (H_2O)	2232	1639	111.2	142.5

graphite-like carbon surfaces, respectively. Each of the states is characterized by the activation energy and pre-exponential factors of the desorption rate constants. The apparent activation energy and heat of sorption were estimated. It was also shown that doping the CDC samples with 0.5 wt% of palladium increased hydrogen uptake by about 2.5 times.

The nondissociative hydrogen uptake for two types of CDC materials was explored in [142]. Multiwalled nanobarrels (MWNBs) prepared by thermochemical treatment of aluminum carbide in chlorine gas [48] and amorphous nanoporous carbon (ANPC) from silicon carbide were used. The specific surface area (BET) of the ANPC was 1000 m^2/g while the MWNB had a density of 0.75 g/cm^3 and S_{BET} = 600 m^2/g. The isotherms of hydrogen adsorption were measured at temperatures of −196, −78, and 27 °C in the pressure range of up to 10 MPa. At the highest hydrogen pressure the maximum differential uptake was determined to 2.9 wt% (27 °C) and 4.1 wt% (−196 °C) for the MWNB and 2.6 wt% (27 °C) and 3.3 wt% (−196 °C) for the ANPC. The total uptake was calculated to be 3.2 wt% (27 °C) and 6.2 wt% (−196 °C) for the MWNB and 2.8 wt% (27 °C) and 4.2 wt% (−196 °C) for the ANPC. Values of hydrogen uptake for the MWNB and ANPC are comparable with other carbonaceous materials (Fig. 12). The larger hydrogen uptake and smaller S_{BET} of MWNB evidently indicated the presence of a significant amount of nanopores in the sample which were inaccessible for nitrogen molecules.

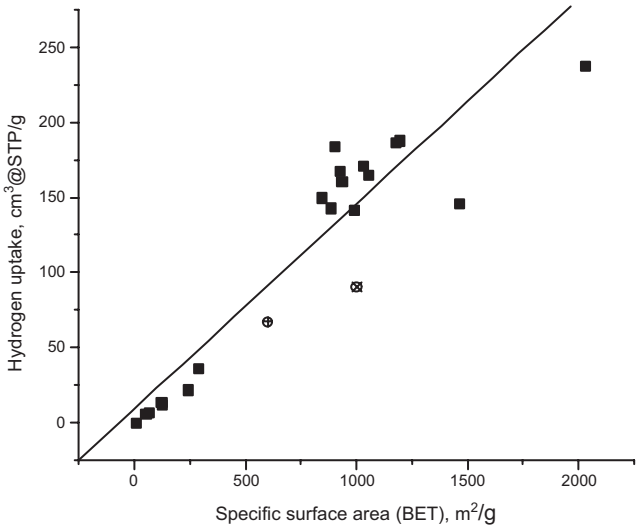


Figure 12. Hydrogen adsorption at −196 °C and 0.1 MPa versus BET surface area for various carbonaceous materials [139]. ⊕—MWNB, ⊗—ANPC [142].

5.3. Lithium-Ion Batteries

Lithium can be inserted reversibly within most carbonaceous materials, and a high amount of reversible insertion is the main requirement for their application as anodes in lithium-ion batteries [143]. A study of lithium diffusion in CDC materials and the phase composition of the produced lithiated samples was recently reported [144, 145]. Bulk carbon specimens prepared according to the technique of Gordeev et al. [32] from carbides of silicon, titanium, and molybdenum were lithiated in the temperature range of 30–200 °C. The concentration profiles of lithium in the carbon specimens were measured using the nuclear reaction $^7Li(p,a)^4He$. The obtained diffusion coefficients were in the range of 10^{-9}–10^{-7} cm^2/s and dependent on the pore size and the diffusion time. Based on these results, a process mechanism is mainly Li diffusion along pore walls and accumulation in pores followed by the formation of metal clusters [144].

The phase composition of the lithiated CDCs was examined by XRD. Intercalation phases LiC_6, LiC_{12}, and LiC_{24} were observed. In some cases, the formation of the undesirable phases of carbonate Li_2CO_3 and carbide Li_2C_2 was detected. Lithiated samples free of carbonate and carbide phases were prepared when the lithium evaporation pulse duration was less than the interval between the impulses.

5.4. Pt-Catalyst on CDC Support

Platinum supported on carbon (Pt/C) is widely used as a catalyst for hydrogenation and oxidation reactions. The porous structure of the carbon support is of primary importance in the catalyst performance. CDC provides the advantages of a narrow pore size distribution and variable pore diameters. Pt/C catalysts are usually prepared by adsorption from a solution of a platinum-containing species on the selected carbon material followed by reduction of the platinum compound to the metal state (see, for example, [146]). This method is applicable to CDC. A novel, one-stage method the for preparation of Pt/C catalysts from SiC was proposed by Ersoy et al. [147, 148]. Silicon carbide ceramics were treated in the Cl_2–H_2–Ar gas mixtures according to the technique described in [10, 63, 64]. Additionally, metallic platinum was present in the reaction zone in the form of foil or wire. In this case, Pt particles were dispersed in the produced material forming interlayers within the carbon layer. The proposed mechanism of the process consists of platinum transport to the SiC/C interface via the gas phase in the form of triplatinum trichloride, after which the Pt_3Cl_3 reacts with Si from the SiC resulting in Pt and $SiCl_4$.

5.5. Tribological Coatings

Metal carbides are widely used in a variety of machining and sliding contact applications. Despite their high corrosion and wear resistance, these materials are not lubricious. A novel process for the formation of a low-friction carbon layer on the SiC surface was recently introduced [2, 65]. Tribological characterization of the silicon carbide ceramics coated using this process was carried out in [149]. Sintered α-SiC and CVD β-SiC specimens were treated in the Cl_2–H_2–Ar mixtures of various composition at temperatures from 700 to 1000 °C and ambient pressure. Tribological tests were performed which measured the friction coefficients and wear rates by means of the standard ball-on-disk method. The produced articles had friction coefficients up to seven times lower than the untreated SiC. The friction coefficients of ~0.1 were almost independent of the type of initial material and the parameters of the ball-on-disk test including contact load and sliding speed. As evidenced by the wear track features, the primary deformation process appears to be plastic flow of the carbon layer, resulting in a smooth surface as the wear process progresses. These coatings can work in dry (nitrogen or vacuum) and humid (air or water) environments, with and without lubricants, and can be used in dynamic seals of water pumps and other tribological applications.

6. CONCLUSIONS

The synthesis of almost all known carbon structures including amorphous and nanocrystalline graphitic carbon, single-wall and multiwall nanotubes, carbon onions, nanocrystalline diamond, and ordered graphite is possible by the extraction of metals from carbides. Since the metal carbide lattice is used as a template and metal is extracted layer by layer, atomic level control can be achieved in the synthesis process and the structure of the carbon can be templated by the carbide structure, with an opportunity for further modification by controlling the temperature, pressure, composition of the environment, and other process variables.

Thermodynamic analysis shows the possibility of carbon synthesis from almost any carbide under appropriate process conditions (temperature, pressure, and etchant). While more than 50% of experimental data have been obtained on SiC, about 20 other carbides have been successfully used to produce CDC.

The linear kinetics of CDC growth in a broad temperature range allows transformation to almost any depth and the synthesis of coatings, powders, and freestanding CDC membranes. This is due to the open nanometer-size porosity in the material. The pore structure of CDC can be tuned very finely by using different carbides, etching agents, and temperature. The average pore size can be controlled with an accuracy of better than 0.1 nm and the pore size distribution produced is narrower than in activated carbons having a similar average pore size which were derived from organic precursors.

Noncatalytic and catalytic synthesis of carbon nanotubes by thermal decomposition of SiC in vacuum has been demonstrated. Catalyst-free nanotube brushes or powders can be produced using this method.

Nanocrystalline diamond can be formed by the extraction of silicon from silicon carbide. It is usually surrounded by amorphous carbon. Continued heat treatment of the nanodiamond at about 1000 °C and higher temperatures results in graphitization and the formation of carbon onions within a few micrometers of the SiC/carbon interface. This onion-containing layer can grow to a thickness of more than 100 μm with uniform structure and properties. The diamond-containing interlayer and the sawlike interface toward the metal carbide explain the excellent adhesion of CDC coatings and the onionlike surface layer is responsible for their excellent tribological properties.

It has been shown that the presence of hydrogen in the environment is not required for diamond synthesis. However, hydrogen can stabilize the diamond nanocrystals in CDC and lead to the growth of thick diamond-structured layers. If no hydrogen is added, diamond nanocrystals transform to graphite, forming carbon onions and other curved graphitic nanostructures.

Low-temperature treatment in halogens produces amorphous carbon with a narrow pore size distribution in the range of 0.5–2 nm. Additionally, mesopores or macropores can be incorporated by varying the starting material (e.g., using porous ceramics) and process parameters. CDC produced from dense ceramics or single crystals has only nanopores and can be used in molecular sieves.

Various approaches to CDC synthesis include decomposition in vacuum, leaching by supercritical water, halogens, and molten salts. CDC structures produced by these techniques have some similarities but also many differences. For example, carbon nanotubes can be produced by high-temperature vacuum decomposition, while amorphous nanoporous carbon is generated by treatment in halogens or supercritical water below 1000 °C.

An important advantage of the CDC process over CVD or PVD for the synthesis of thin carbon coatings is that the transformation is conformal and does not lead to a change of the carbide sample's size or shape. This may be of particular importance in tribological coatings or in nano/microtechnology applications when high surface quality and nanometer range accuracy must be achieved. Powders, fibers, and carbide components of any shape can be coated with carbon using the CDC process.

CDC synthesis can be environmentally friendly even when halogens are used for metal extraction. This process allows the utilization of metal halogenides generated as a by-product of carbon synthesis. Alternatively, the halogenides can be decomposed in the high-temperature zone of the furnace to produce a metal and halogen, which can be returned to the reactor creating a closed-cycle process. The use of low-cost furnaces and moderate temperatures in halogen-assisted synthesis allows the large-scale production of nanostructured CDC.

Potential applications of CDC include supercapacitors, lithium ion batteries, molecular sieves, gas storage (hydrogen, methane and other), gas separation, filtration of liquids and gases, catalyst support, and adsorbents for soil and air decontamination, gas chromatography, and other purposes. CDC films can be used as tribological coatings in dynamic seals and other applications involving sliding or rolling friction. They have a low friction coefficient of 0.02 to 0.2 and,

unlike other carbon materials, work well in both wet and dry environments. Specific structures produced by this process can find other applications in nanoelectronics and electrochemistry, or as ultralight structural or multifunctional materials.

GLOSSARY

Adsorption isotherm An adsorption isotherm for a single gaseous adsorptive on a solid is the function which relates at constant temperature the amount of substance adsorbed at equilibrium to the pressure (or concentration) of the adsorptive in the gas phase.

Amorphous carbon A carbon material without long-range crystalline order. Short-range order exists, but with deviations of the interatomic distances and/or interbonding angles with respect to the graphite lattice as well as to the diamond lattice.

Carbide-derived carbon (CDC) Carbon material produced by selective extraction of metal atoms from a carbide crystal lattice by halogens, supercritical water, oxygen at a low partial pressure, or other etchants, or in vacuum.

Carbon nanotube A nanoscale (1–100 nm) carbon structure produced by rolling graphene sheet(s) into a tube.

Graphitization A solid-state transformation of thermodynamically unstable nongraphitic carbon into graphite by means of heat treatment.

Hydrothermal synthesis Synthesis of material using supercritical water.

Lithium-ion battery An electrochemical battery based on lithium metal as anode.

Nanodiamond A diamond with particle size of 1–10 nm produced by shock synthesis, chemical vapor deposition, or a CDC process.

Nanoporous carbon A porous carbon material, which is considered to have a major part of its porosity in pores of less than 2 nm width and to exhibit apparent surface areas usually of 200 to 2000 m^2/g. Also called microporous carbon following IUPAC classification.

Physical adsorption Adsorption in which the forces involved are intermolecular forces (van der Waals forces) of the same kind as those responsible for the imperfection of real gases and the condensation vapors, and which do not involve a significant change in the electronic orbital patterns of the species involved.

Specific surface area When the area of the interface between two phases is proportional to the mass of one of the phases (e.g., for a solid adsorbent, for an emulsion, or for an aerosol), the specific surface area can be defined as the surface area can be defined as the surface area divided by the mass of the relevant phase.

Supercapacitor Electrochemical double layer capacitor. The electric energy is stored in the electric double layer from charge separation which is electrostatic force between an ionically conducting electrolyte and a conducting electrode. The ions displaced in forming the double layers are transferred between the electrodes by diffusion through the electrolyte.

Tribology The branch of engineering that deals with the interaction of surfaces in relative motion (as in bearings or gears): their design, friction, wear, and lubrication.

ACKNOWLEDGMENTS

This work was supported by DARPA via ONR contract. Y.G. would like to acknowledge productive collaboration with Professor M. J. McNallan (UIC) as well as assistance from his students and postdoctoral associates H. Ye, B. Carrol, and L. Chen at Drexel University and S. Welz at the University of Illinois at Chicago, who worked with him on development of new CDC materials.

REFERENCES

1. Y. Gogotsi, V. Kamyshenko, V. Shevchenko, S. Welz, D. A. Ersoy, and M. J. McNallan, in "Proc. NATO ASI on Functional Gradient Materials and Surface Layers Prepared by Fine Particles Technology" (M.-I. Baraton and I. Uvarova, Eds.), p. 239. Kluwer Academic, Dordrecht, 2001.
2. D. A. Ersoy, M. J. McNallan, and Y. Gogotsi, *Mater. Res. Innovat.* 5, 55 (2001).
3. W. A. Mohun, in "Proc. 4th Biennial Conf. Carbon," p. 443. Pergamon, Oxford, 1959.
4. M. Kusunoki, M. Rokkaku, and T. Suzuki, *Appl. Phys. Lett.* 71, 2620 (1997).
5. Y. Gogotsi, S. Welz, D. A. Ersoy, and M. J. McNallan, *Nature* 411, 283 (2001).
6. O. Hutchins, U.S. Patent 1, 271, 713, 1918.
7. J. N. Andersen, U.S. Patent 2, 739, 041, 1956.
8. G. F. Kirillova, G. A. Meerson, and A. N. Zelikman, *Izv. Tsvetnaya. Metall.* 90 (1960).
9. V. P. Orekhov, G. V. Seryakov, A. N. Zelikman, T. M. Starobina, T. I. Kahzanova, K. V. Petrova, and P. M. Sverchkov, *J. Appl. Chem. USSR* 42, 230 (1969).
10. Y. G. Gogotsi, I.-D. Jeon, and M. J. McNallan, *J. Mater. Chem.* 7, 1841 (1997).
11. M. J. McNallan, Y. G. Gogotsi, and I. Jeon, in "Proceedings of the Ninth International Conference on High Temperature Materials Chemistry" (K. E. Spear, Ed.), Vol. 97-39, p. 529. The Electrochemical Society, Pennington, NJ, 1997.
12. W. A. Mohun, U.S. Patent 3, 066, 099, 1962.
13. Nillok Chemicals, Great Britain Patent 971943, 1964.
14. H. P. Boehm and H. H. Warnecke, in "Proc. 12th Biennial Conf. on Carbon," p. 149. Pergamon, Oxford, 1975.
15. N. F. Fedorov, G. K. Ivakhnyuk, V. V. Tetenov, and G. V. Matyukhin, *J. Appl. Chem. USSR* 54, 1239 (1981).
16. N. F. Fedorov, G. K. Ivakhnyuk, and D. N. Gavrilov, *Zh. Prikladnoi Khim.* 55, 46 (1982).
17. N. F. Fedorov, G. K. Ivakhnyuk, and D. N. Gavrilov, *Zh. Prikladnoi Khim.* 55, 272 (1982).
18. O. E. Babkin, G. K. Ivakhnyuk, and N. F. Fedorov, *Zh. Prikladnoi Khim.* 57, 504 (1984).
19. K. S. W. Sing, D. H. Everett, R. A. V. Haul, L. Moscou, R. A. Pierotti, J. Rouquerol, and T. Siemieniewska, *Pure Appl. Chem.* 57, 603 (1985).
20. J. Leis, A. Perkson, M. Arulepp, P. Nigu, and G. Svensson, *Carbon* 40, 1559 (2002).
21. M. Jacob, U. Palmqvist, P. Alberius-Henning, and T. Ekström, *Mater. Res. Soc. Symp. Proc.* 593, 87 (2000).
22. J. Leis, A. Perkson, M. Arulepp, M. Kaarik, and G. Svensson, *Carbon* 39, 2043 (2001).
23. N. F. Fedorov, *Russ. Chem. J.* 39, 73 (1995).

24. Y. A. Kukushkina, R. G. Avarbe, and V. V. Sokolov, *Russ. J. Appl. Chem.* 69, 620 (1996).
25. Y. A. Kukushkina, R. G. Avarbe, V. V. Sokolov, and T. V. Mazaeva, *Russ. J. Appl. Chem.* 69, 623 (1996).
26. N. A. Yanul', Y. A. Kukushkina, V. V. Sokolov, S. K. Gordeev, A. E. Kravchik, and Y. E. Kirsh, *Russ. J. Appl. Chem.* 72, 2159 (1999).
27. Y. E. Kirsh, N. P. Bessonova, N. A. Yanul', S. K. Gordeev, and S. F. Timashev, *Russ. J. Phys. Chem.* 74, 540 (2000).
28. S. K. Gordeev, in "NATO ASI on Nanostructured Carbon for Advanced Applications" (G. Benedek, P. Milani, and V. G. Ralchenko, Eds.), p. 71. Kluwer Academic, Dordrecht, 2001.
29. S. K. Gordeev and A. V. Vartanova, *Zh. Prikladnoi Khim.* 64, 1178 (1991).
30. S. K. Gordeev and A. V. Vartanova, *Zh. Prikladnoi Khim.* 67, 1080 (1994).
31. S. K. Gordeev and A. V. Vartanova, *Zh. Prikladnoi Khim.* 67, 1375 (1994).
32. S. K. Gordeev, R. G. Avarbz, A. E. Kravtjik, J. A. Kukusjkina, V. V. Sokolov, T. V. Mazaeva, and A. Grechinskaya, PCT Patent WO 98/54111, 1998.
33. G. K. Ivakhnyuk, O. E. Babkin, and N. F. Fedorov, *Russ. J. Phys. Chem.* 67, 1849 (1993).
34. Y. A. Kukushkina, R. G. Avarbe, V. V. Sokolov, and A. E. Kravchik, *Russ. J. Appl. Chem.* 72, 2144 (1999).
35. S. A. Shevchenko, G. M. Romanova, Y. A. Kukushkina, and R. G. Avarbe, *Russ. J. Appl. Chem.* 72, 2153 (1999).
36. V. B. Moshkovskii, R. G. Avarbe, Y. A. Kukushkina, and V. V. Sokolov, *Russ. J. Appl. Chem.* 72, 2155 (1999).
37. F. Euler and E. R. Czerlinsky, Silicon carbide. A high temperature semiconductor, in "Proceedings of the Conference" (J. R. O'Connor and J. Smiltens, Eds.), p. 155. Pergamon, Oxford, 1960.
38. L. F. Lowe, H. D. Thompson, and J. P. Cali, Silicon carbide. A high temperature semiconductor, in "Proceedings of the Conference" (J. R. O'Connor and J. Smiltens, Eds.), p. 221. Pergamon, Oxford, 1960.
39. H. P. Boehm, *Carbon* 35, 581 (1997).
40. G. R. Millward and D. A. Jefferson, in "Chemistry and Physics of Carbon" (P. L. Walker, Jr. and P. A. Thrower, Eds.), Vol. 14, p. 1. Dekker, New York, 1978.
41. R. N. Kyutt, E. A. Smorgonskaya, A. M. Danishevskii, S. K. Gordeev, and A. V. Grechinskaya, *Phys. Solid State* 41, 808 (1999).
42. R. N. Kyutt, E. A. Smorgonskaya, A. M. Danishevski, S. K. Gordeev, and A. V. Grechinskaya, *Phys. Solid State* 41, 1359 (1999).
43. E. A. Smorgonskaya, R. N. Kyutt, S. K. Gordeev, A. V. Grechinskaya, Y. A. Kukushkina, and A. M. Danishevskii, *Phys. Solid State* 42, 1176 (2000).
44. E. A. Smorgonskaya, R. N. Kyutt, A. V. Shchukarev, S. K. Gordeev, and A. V. Grechinskaya, *Semiconductors* 35, 661 (2001).
45. V. V. Popov, S. K. Gordeev, A. V. Grechinskaya, and A. M. Danishevskii, *Phys. Solid State* 44, 789 (2002).
46. S. K. Gordeev, S. A. Kukushkin, A. V. Osipov, and Y. V. Pavlov, *Phys. Solid State* 42, 2314 (2000).
47. E. Smorgonskaya, R. Kyutt, A. Danishevskii, C. Jardin, R. Meaudre, O. Marty, S. Gordeev, and A. Grechinskaya, *J. Non-Cryst. Solids* 299–302 (2002).
48. J. Zheng, T. C. Eckström, S. K. Gordeev, and M. Jacob, *J. Mater. Chem.* 10, 1039 (2000).
49. S. Welz, Y. Gogotsi, and M. J. McNallan, *J. Appl. Phys.* 93, 4207 (2003).
50. T. Ekström, M. Jacob, J. Zheng, P. Berius-Henning, U. Palmqvist, J. Leis, and A. Perkson, PCT Patent WO 01/16023, 2001.
51. A. M. Danishevskii, E. A. Smorgonskaya, S. K. Gordeev, and G. A. V, *Phys. Solid State* 43, 137 (2001).
52. M. J. McNallan, S. Y. Ip, S. Saam, and W. W. Liang, in "High Temperature Materials Chemistry—III" (Z. A. Munir and D. Cubiciotti, Eds.), p. 328. The Electrochemical Society, Pennington, NJ, 1986.
53. S. Y. Ip, M. J. McNallan, and M. E. Schreiner, *Ceram. Trans.* 2, 289 (1989).
54. M. J. McNallan, S. Y. Ip, S. Y. Lee, and C. Park, *Ceram. Trans.* 10, 309 (1990).
55. D. S. Park, M. J. McNallan, C. Park, and W. W. Liang, *J. Am. Ceram. Soc.* 73, 1323 (1990).
56. J. E. Marra, E. R. Kreidler, N. S. Jacobson, and D. S. Fox, *J. Am. Ceram. Soc.* 71, 1067 (1988).
57. J. E. Marra, E. R. Kreidler, N. S. Jacobson, and D. S. Fox, *Ceram. Trans.* 2, 275 (1989).
58. J. E. Marra, E. R. Kreidler, N. S. Jacobson, and D. S. Fox, *J. Electrochem. Soc.* 135, 1571 (1988).
59. M. Balooch and D. R. Olander, *Surf. Sci.* 261, 321 (1992).
60. W.-S. Pan and A. J. Steckl, *J. Electrochem. Soc.* 137, 212 (1990).
61. P. H. Yih and A. J. Steckl, *J. Electrochem. Soc.* 140, 1813 (1993).
62. P. H. Yih and A. J. Steckl, *J. Electrochem. Soc.* 142, 2853 (1995).
63. I. D. Jeon, M. J. McNallan, and Y. G. Gogotsi, in "Proceedings of The Symposium on Fundamental Aspects of High Temperature Corrosion" (D. A. Shores, R. A. Rapp, and P. Y. Hou, Eds.), Vol. 96-26, p. 256. The Electrochemical Society, Pennington, NJ, 1996.
64. Y. G. Gogotsi, in "Advanced Multilayered and Fibre-Reinforced Composites" (Y. M. Haddad, Ed.), p. 217. Kluwer Academic, Dordrecht, 1998.
65. D. A. Ersoy, M. J. McNallan, and Y. Gogotsi, in "Proceedings of the Symposium on High Temperature Corrosion and Materials Chemistry" (P. Y. Hou, M. J. McNallan, R. Oltra, E. J. Opila, and D. A. Shores, Eds.), Vol. 98-9, p. 324. The Electrochemical Society, Pennington, NJ, 1998.
66. M. McNallan, Y. Gogotsi, and I. Jeon, in "Tribology Issues and Opportunities in MEMS" (B. Bhushan, Ed.), p. 559. Kluwer Academic, Dordrecht, 1998.
67. Y. G. Gogotsi, S. Welz, J. Daghfal, M. J. McNallan, I.-D. Jeon, K. G. Nickel, and T. Kraft, *Ceram. Eng. Sci. Proc.* 19, 87 (1998).
68. M. J. McNallan, Y. Gogotsi, and D. A. Ersoy, in "Elevated Temperature Coatings: Science and Technology III" (J. M. Hampikian and N. B. Dahotre, Eds.), p. 351. The Minerals, Metals & Materials Society, 1999.
69. V. V. Kamyshenko, V. I. Shevchenko, Y. G. Gogotsi, and V. V. Kartuzov, in "Mathematical Models and Computing Experiment in Materials Science" (B. A. Galanov, Ed.), Vol. 5, p. 65. Frantsevich Institute for Problems of Materials Science, Kiev, 2001.
70. O. E. Babkin, G. K. Ivakhnyuk, Y. N. Lukin, and N. F. Fedorov, *Zh. Prikladnoi Khim.* 57, 1719 (1984).
71. T. Y. Kosolapova, "Carbides. Properties, Production, and Applications." Plenum, New York/London, 1971.
72. D. Osetzky, *Carbon* 12, 517 (1974).
73. N. F. Fedorov, G. K. Ivakhnyuk, and V. V. Samonin, *J. Appl. Chem. USSR* 54, 2253 (1981).
74. G. K. Ivakhnyuk, V. V. Samonin, N. F. Fedorov, V. A. Vladimirov, L. V. Stepanova, and O. M. Kas'yanova, *Zh. Prikladnoi Khim.* 60, 852 (1987).
75. G. K. Ivakhnyuk, V. V. Samonin, and N. F. Fedorov, *Zh. Prikladnoi Khim.* 60, 1413 (1987).
76. V. V. Samonin, G. K. Ivakhnyuk, and N. F. Fedorov, *Zh. Prikladnoi Khim.* 60, 2357 (1987).
77. N. F. Fedorov and V. V. Samonin, *Russ. J. Appl. Chem.* 71, 584 (1998).
78. N. F. Fedorov and V. V. Samonin, *Russ. J. Appl. Chem.* 71, 795 (1998).
79. N. S. Jacobson, Y. G. Gogotsi, and M. Yoshimura, *J. Mater. Chem.* 5, 595 (1995).

80. S. Shimada, M. Inagaki, and K. Matsui, *J. Am. Ceram. Soc.* 75, 2671 (1992).
81. C. B. Bargeron and R. C. Benson, *J. Am. Ceram. Soc.* 76, 1040 (1993).
82. S. Shimada, K. Nakajima, and M. Inagaki, *J. Am. Ceram. Soc.* 80, 1749 (1997).
83. S. Shimada, F. Yunazar, and S. Otani, *J. Am. Ceram. Soc.* 83, 721 (2000).
84. S. Shimada, *Mater. Sci. Forum* 369–372, 377 (2001).
85. S. Shimada, *Solid State Ionics* 149, 319 (2002).
86. S. Shimada and T. Ishii, *J. Am. Ceram. Soc.* 73, 2804 (1990).
87. S. Shimada, M. Nishisako, M. Inagaki, and K. Yamamoto, *J. Am. Ceram. Soc.* 78, 41 (1995).
88. S. Shimada, *Solid State Ionics* 101–103, 749 (1997).
89. S. Shimada, M. Yoshimatsu, M. Inagaki, and S. Otani, *Carbon* 36, 1125 (1998).
90. Y. G. Gogotsi and M. Yoshimura, *Nature* 367, 628 (1994).
91. Y. G. Gogotsi and M. Yoshimura, *J. Am. Ceram. Soc.* 78, 1439 (1995).
92. Y. Gogotsi and M. Yoshimura, *J. Mater. Sci. Lett.* 14, 755 (1995).
93. Y. G. Gogotsi, Y. Tanabe, E. Yasuda, and M. Yoshimura, in "Advanced Materials '93, I/A: Ceramics, Powders, Corrosion and Advanced Processing" (N. Mizutani et al., Ed.). Elsevier Science, Amsterdam, 1994.
94. Y. Gogotsi, in "Korrosion und Verschleiß von keramischen Werkstoffen" (R. Telle and P. Quirmbach, Eds.), p. 114. Deutsche Keramische Gesellschaft, Aachen, 1994.
95. Y. G. Gogotsi, M. Yoshimura, M. Kakihana, Y. Kanno, and M. Shibuya, in "Ceramic Processing Science and Technology" (H. Hausner, G. L. Messing, and S.-I. Hirano, Eds.), p. 243. Am. Ceram. Soc., Westerville, OH, 1995.
96. T. Kraft, K. G. Nickel, and Y. G. Gogotsi, *J. Mater. Sci.* 33, 4357 (1998).
97. T. Kraft and K. G. Nickel, *J. Mater. Chem.* 10, 671 (2000).
98. Y. G. Gogotsi, K. G. Nickel, and P. Kofstad, *J. Mater. Chem.* 5, 2313 (1995).
99. Y. G. Gogotsi, P. Kofstad, M. Yoshimura, and K. G. Nickel, *Diamond Relat. Mater.* 5, 151 (1996).
100. Y. G. Gogotsi and K. G. Nickel, *Ceram. Eng. Sci. Proc.* 18, 747 (1997).
101. R. Roy, D. Ravichandran, A. Badzian, and E. Breval, *Diamond Relat. Mater.* 5, 973 (1996).
102. K. G. Nickel, T. Kraft, and Y. G. Gogotsi, in "Handbook of Ceramic Hard Materials" (R. Riedel, Ed.), p. 374. Wiley–VCH, Weinheim, 2000.
103. A. A. Giardini and J. E. Tydings, *Am. Mineral.* 47 (1962).
104. S. Naka, A. Tsuzuki, and S.-I. Hirano, *J. Mater. Sci.* 19, 259 (1984).
105. U. Gerlach, H. Hennig, H. Oppermann, G. Risse, and E. Wolf, East Germany Patent DD 257375, 1988.
106. E. Wolf, H. Oppermann, H. Hennig, U. Gerlach, G. Grosse, and G. Risse, East Germany Patent, DD 259147, 1988.
107. S. Hong and M. Wakatsuki, *J. Mater. Sci. Lett.* 12, 283 (1993).
108. S. M. Hong, W. Li, X. P. Jia, and M. Wakatsuki, *Diamond Relat. Mater.* 2, 508 (1993).
109. L. Gou, S. Hong, and Q. Gou, *J. Mater. Sci.* 30, 5687 (1995).
110. X. J. Luo, Q. Liu, L. Ding, and B. C. Luo, *J. Mater. Sci. Lett.* 16, 1005 (1997).
111. L. W. Yin, M. S. Li, Y. X. Liu, D. S. Sun, F. Z. Li, G. L. Geng, and Z. Y. Hao, *Appl. Phys. A* 72, 487 (2001).
112. L. W. Yin, M. S. Li, Y. X. Liu, B. Xu, P. Liu, and Z. Y. Hao, *Chem. Phys. Lett.* 363, 211 (2002).
113. W. F. Knippenberg, *Philips Res. Rep.* 18, 161 (1963).
114. D. V. Badami, *Nature* (*London*) 193, 569 (1962).
115. D. V. Badami, *Carbon* 3, 53 (1965).
116. J. J. Comer, *J. Appl. Cryst.* 4, 12 (1971).
117. A. J. van Bommel, J. E. Crombeen, and A. van Tooren, *Surf. Sci.* 48, 463 (1975).
118. F. Meyer and G. J. Loyen, *Acta Electron.* 18, 33 (1975).
119. S. Iijima, *J. Solid State Chem.* 42, 101 (1982).
120. K. Motzfeldt and M. Steinmo, in "Proceedings of the Ninth International Conference on High Temperature Materials Chemistry" (K. E. Spear, Ed.), Vol. 97-39, p. 523. The Electrochemical Society, Pennington, NJ, 1997.
121. N. I. Voronin, V. L. Kuznetsova, and R. I. Bresker, in "Silicon Carbide" (I. N. Frantsevich, Ed.), p. 81. Consultants Bureau, New York/London, 1970.
122. L. M. Foster, G. Long, and H. C. Stumpf, *Am. Mineral.* 43, 285 (1958).
123. H. Takikawa, R. Miyano, M. Yatsuki, and T. Sakakibara, *Jpn. J. Appl. Phys.* 37, L187 (1998).
124. M. Kusunoki, J. Shibata, M. Rokkaku, and T. Hirayama, *Jpn. J. Appl. Phys.* 37, L605 (1998).
125. M. Kusunoki, T. Suzuki, K. Kaneko, and M. Ito, *Philos. Mag. Lett.* 79, 153 (1999).
126. H. Watanabe, Y. Hisada, S. Mukainakano, and N. Tanaka, *J. Microsc.* 203, 40 (2001).
127. M. Kusunoki, T. Suzuki, T. Hirayama, J. Shibata, and K. Kaneko, *Appl. Phys. Lett.* 77, 531 (2000).
128. M. Kusunoki, T. Suzuki, T. Hirayama, and N. Shibata, *Physica B* 323, 296 (2002).
129. T. Shimizu, Y. Ishikawa, M. Kusunoki, T. Nagano, and N. Shibata, *Jpn. J. Appl. Phys.* 39, L1057 (2000).
130. V. Derycke, R. Martel, M. Radosavljević, F. M. Ross, and P. Avouris, *Nano Lett.* 2, 1043 (2002).
131. M. Kusunoki, T. Suzuki, C. Honjo, T. Hirayama, and N. Shibata, *Chem. Phys. Lett.* 366, 458 (2002).
132. K. Tsugawa, A. Goto, and Y. Koga, in "Proc. of 6th ADC/FCT Conf.," 2001, p. 800.
133. A. Burke, *J. Power Sources* 91, 37 (2000).
134. R. G. Avarbz, A. V. Vartanova, S. K. Gordeev, S. G. Zjukov, B. A. Zelenov, A. E. Kravtjik, V. P. Kuznetsov, J. A. Kukusjkina, T. V. Mazaeva, O. S. Pankina, and V. V. Sokolov, PCT Patent WO 97/20333, 1997.
135. R. G. Avarbz, A. V. Vartanova, S. K. Gordeev, S. G. Zjukov, B. A. Zelenov, A. E. Kravtjik, V. P. Kuznetsov, J. A. Kukusjkina, T. V. Mazaeva, O. S. Pankina, and V. V. Sokolov, U.S. Patent 5, 876, 787, 1999.
136. R. G. Avarbz, A. V. Vartanova, S. K. Gordeev, S. G. Zjukov, B. A. Zelenov, A. E. Kravtjik, V. P. Kuznetsov, J. A. Kukusjkina, T. V. Mazaeva, O. S. Pankina, and V. V. Sokolov, U.S. Patent 6, 110, 335, 2000.
137. Y. A. Maletin, N. G. Strizhakova, V. G. Izotov, A. A. Mironova, S. G. Kozachkov, V. G. Danilin, S. N. Podmogilny, M. Arulepp, J. A. Kukusjkina, A. E. Kravtjik, V. V. Sokolov, A. Perkson, J. Leis, J. Zheng, S. K. Gordeev, J. Y. Kolotilova, J. Cederstrom, and C. L. Wallace, PCT Patent WO 02/39468, 2002.
138. L. Schlapbach and A. Züttel, *Nature* 414, 353 (2001).
139. M. G. Nijkamp, J. E. M. J. Raaymakers, A. J. van Dillen, and K. P. de Jong, *Appl. Phys. A* 72, 619 (2001).
140. E. A. Evard, A. P. Voit, S. K. Gordeev, and I. E. Gabis, *Mater. Sci.* 36, 499 (2000).
141. I. E. Gabis, E. A. Evard, S. K. Gordeev, and T. Ekström, in "Hydrogen Materials Science and Chemistry of Metal Hydrides" (M. D. Hampton, Ed.), p. 383. Kluwer Academic, Dordrecht, 2002.
142. E. Johansson, B. Hjorvarsson, T. Ekström, and M. Jacob, *J. Alloys Compounds* 330–332, 670 (2002).
143. S. Flandrois and B. Simon, *Carbon* 37, 165 (1999).
144. I. M. Kotina, V. M. Lebedev, A. G. Ilves, G. V. Patsekina, L. M. Tuhkonen, S. K. Gordeev, M. A. Yagovkina, and T. Ekström, *J. Non-Cryst. Solids* 299–302, 815 (2002).

145. I. M. Kotina, V. M. Lebedev, A. G. Ilves, G. V. Patsekina, L. M. Tuhkonen, S. K. Gordeev, M. A. Yagovkina, and T. Ekström, *J. Non-Cryst. Solids* 299–302, 820 (2002).
146. E. Antolini, F. Gardellini, E. Giacometti, and G. Squadrito, *J. Mater. Sci.* 37, 133 (2002).
147. D. A. Ersoy, M. J. McNallan, and Y. Gogotsi, in "High Temerature Corrosion and Materials Chemistry: Per Kofstad Memorial Symposium" (P. Y. Hou, T. Maruyama, M. J. McNallan, T. Narita, E. J. Opila, and D. A. Shores, Eds.), Vol. 99-38, p. 456. The Electrochemical Society, Pennington, NJ, 1999.
148. D. A. Ersoy, M. J. McNallan, and Y. Gogotsi, *J. Electrochem. Soc.* 148, C774 (2001).
149. D. A. Ersoy, M. J. McNallan, Y. Gogotsi, and A. Erdemir, *Tribol. Trans.* 43, 809 (2000).
150. Y. Gogotsi, A. Nikitin, H. Ye, W. Zhou, J. E. Fischer, B. Yi, H. C. Foley, and M. W. Barsoum, *Nature Materials* 2, 591 (2003).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanostructured Barium Strontium Titanate Thin Films

Bonnie D. Riehl, Guru Subramanyam

University of Dayton, Dayton, Ohio, USA

Rand R. Biggers, Angela Campbell

Air Force Research Laboratory/Materials and Manufacturing Directorate, Wright-Patterson AFB, Ohio, USA

CONTENTS

1. INTRODUCTION

The interest in ferroelectrics arises because a ferroelectric material exhibits spontaneous polarization. A ferroelectric crystal consists of positive and negative ions. For a certain temperature range, the positive and negative ions become displaced. This results in a net dipole moment. The orientation can be reversed with the application of an electric field. Entry into the spontaneous polarization temperature range requires the ferroelectric crystals to undergo a phase transition, which involves structural changes. The two phases, with and without a spontaneous dipole moment, are referred to as the ferroelectric and paraelectric respectively [1]. As the temperature decreases from above the Curie temperature, T_c, a structural phase change occurs and the crystal changes from paraelectric to ferroelectric. Since the potential energy in the ferroelectric phase of the two polarization orientations has the same energetic minimum, both polarizations are equally favored. Interest in ferroelectrics has grown rapidly because the high relative dielectric constant of a ferroelectric material can be used in dynamic random access memories (DRAM). Its capacity for being polarized in opposite directions makes it a candidate for nonvolatile random access memories. Ferroelectric thin films are also frequency agile due to the nonlinear dc electric-field dependence of their relative dielectric constant. This characteristic can be used for frequency and phase agile microwave circuits. Two popular ferroelectric thin films for frequency agile circuits are strontium titanate ($SrTiO_3$) and barium strontium titanate ($BaSrTiO_3$).

One aspect of studying particle size effects in ferroelectrics is to determine the ultimate level to which a device fabricated from these materials can be miniaturized without degrading the desirable characteristics of the material. The experimental study of size effects in ferroelectrics has gained momentum because of the advances in synthesis and material characterization techniques [2–10]. These advances provide the means for the study of size effects in nanostructured ferroelectrics.

1.1. Ferroelectric Materials

1.1.1. Dielectrics/Ferroelectrics

A dielectric material is one that is electrically insulating (nonmetallic) and exhibits, or can be made to exhibit, an electric dipole structure. There is a separation of positive and negative electrically charged particles at either a molecular or atomic level. Dielectric polarization can be visualized as the action of dipole chains, which form under the influence of an applied field. The countercharges are bound by their free ends when terminated on metal surfaces [1].

The polarization $\mathbf{P}$ ($\mathrm{C\,m^{-2}}$) induced in a dielectric by an applied electric field vector $\mathbf{E}$ ($\mathrm{V\,m^{-1}}$) is $\chi\varepsilon_0\,\mathbf{E}$, where χ ($\mathrm{F\,m^{-1}}$) is the dielectric susceptibility of the material. The total surface charge density induced by an applied electric field, $\mathbf{D}$ ($\mathrm{C\,m^{-2}}$), is given by $\mathbf{D} = \varepsilon_0\mathbf{E} + \mathbf{P}$, $\mathbf{D}$ ($\mathrm{C\,m^{-2}}$) is the dielectric displacement vector, and ε_0, given by

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (575–594)

8.854×10^{-12} F m^{-1}, is the dielectric permittivity of vacuum. Complex permittivity $\varepsilon^* = \varepsilon' - i\varepsilon''$ is used to account for migration of charge carriers as well as any other energy-consuming process that may occur in the material. The loss tangent is defined as $\tan\delta = \varepsilon''/\varepsilon'$ and the quality factor $Q = 1/\tan\delta$ [1] (Fig. 1).

Some polar materials, called pyroelectric, display an electric dipole moment even in the absence of an external electric field. The polarization associated with a spontaneously formed dipole moment is called spontaneous polarization, P_s. Piezoelectric materials can be polarized by the application of an electric field as well as by the addition of a mechanical stress. Ferroelectric materials are both pyroelectric and piezoelectric [1].

Ferroelectric materials are polar and possess at least two equilibrium orientations of the spontaneous polarization vector in the absence of an applied electric field. The spontaneous polarization vector may be switched between these orientations by an applied electric field. Most ferroelectrics undergo a structural phase transition from a high temperature nonferroelectric (called paraelectric) phase to a low-temperature ferroelectric phase. T_c, the Curie temperature, is the temperature of the phase transition. At temperatures above T_c the dielectric permittivity falls off with temperature and follows the Curie–Weiss law [1]. For the ferroelectrics (example, $BaTiO_3$) that undergo several phase transitions into successive ferroelectric phases, only the first ferroelectric phase is called the Curie point. The ferroelectric phase transition causes anomalies in the dielectric, elastic, and thermal properties. The transition is also accompanied with changes in the dimensions of the crystal unit cell. There is a spontaneous strain, x_s, that is associated with the phase transition. It is responsible for the difference in the dimensions of the ferroelectric and paraelectric unit cells [11]. A change that can occur in a ferroelectric material is the transformation from a paraelectric cubic into a ferroelectric tetragonal phase.

In the ferroelectric phase, a crystal is spontaneously strained with $a_t \leq a_c < c_t$ where a_t and a_c are the a-axes of the tetragonal and cubic unit cell. Spontaneous polarization can arise along any of these axes with equal probability when the crystal is cooled through the ferroelectric phase transition temperature. Polarization will develop in directions dependent upon the electrical and mechanical boundary conditions imposed on the material. Regions of uniformly oriented spontaneous polarization are called ferroelectric domains. Regions between two domains are called domain walls. Domain walls oriented differently in response to the spontaneous polarization vector are called ferroelectric domain walls, and those which differ in orientation to the spontaneous strain tensor are called ferroelastic domain walls (Fig. 2). For a detailed discussion see [1].

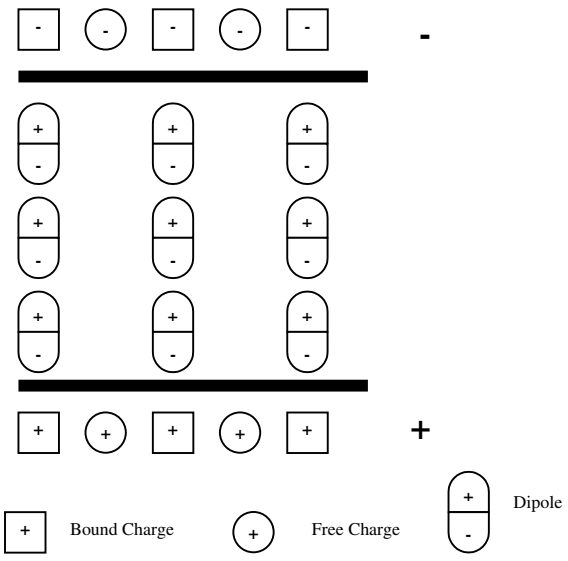


Figure 1. Dipole chains for a ferroelectric material.

The most important characteristic of ferroelectric materials is the reversal of polarization by an electric field. A consequence of the domain-wall switching in a ferroelectric materials is the ferroelectric hysteresis loop, as shown in Figure 3. At small values of the applied electric field, the polarization increases linearly. In this region the field is not strong enough to switch domains with the unfavorable direction of polarization. As the field is increased, the polarization of domains with an unfavorable direction of polarization will start to switch in the direction of the field. Zero polarization requires reversing the field.

Further increasing of the field in the negative direction will continue the realignment of dipoles until saturation. The field is then reduced to zero and reversed to complete the cycle [1]. The value of polarization at $E = 0$ is called the remnant polarization, P_r. E_c is called the coercive field, where the polarization is zero. The spontaneous polarization P_s is the intercept of the polarization axis. The coercive field, spontaneous and remnant polarization, and shape of the hysteresis loop may be affected by the film thickness, the presence of charged defects, mechanical stresses, preparation conditions, and thermal treatment [12]. The slope of the P vs E characteristics is a direct measure of ε_r of the ferroelectric material. The nonlinear P–E characteristics result in nonlinear electric field dependence of the ε_r.

When a periodically varying electric field E is applied to a ferroelectric material a hysteresis loop is obtained. This dynamic hysteresis is the representative characteristic of a ferroelectric. The hysteresis area, A, represents the energy dissipation within one period of domain reversal. The area is the characteristic parameter used to scale the dynamics of nonequilibrium first-order phase transitions in ferroelectrics. The kinetic nucleation and growth of new domains can be

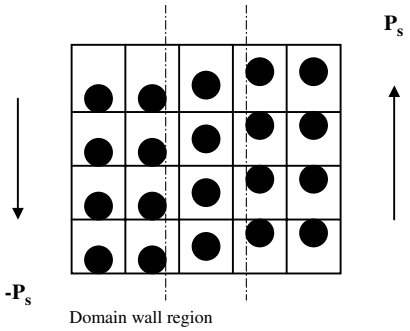


Figure 2. Ferroelectric domain regions in response to spontaneous polarization.

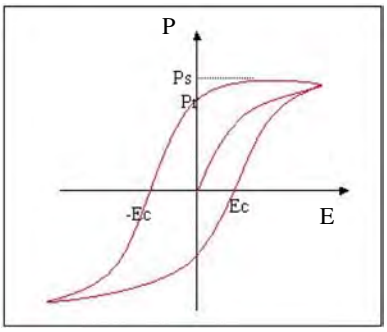


Figure 3. Hysteresis curve for a ferroelectric material under an applied electric field.

evaluated from the details of the hysteresis loop. The remnant polarization P_r and coercive field E_c can be determined from the hysteresis loop [13].

1.2. Applications of Ferroelectric Materials

1.2.1. Memory

Ferroelectric memory possesses advantages such as nonvolatility, low power consumption, high endurance, and high speed writing. The manufacturing processes for ferroelectrics are more compatible with the standard complementary metal–oxide–semiconductor (CMOS) process than traditional nonvolatile memory process. They do not require high voltage operation and the ferroelectric process does not affect characteristics of CMOS devices used in logic cells, analog cells, and core cells [14].

The importance of memory to electronic data processing drives continuous efforts to develop improved memory devices. Obstacles to very large scale integrated fabrication arise because the area of memory storage capacitance must be maintained while the device dimensions are reduced. This provides impetus to the adoption of complicated structures like the trench or the stack causing the number of processing steps to increase. The use of high dielectric constant materials has opened another avenue of DRAM development. Materials being studied in an effort to fabricate a film with a higher dielectric constant are $SrTiO_3$ and $(Ba,Sr)TiO_3$ [15].

1.2.2. Microwave

High relative dielectric constant (ε_r) ferroelectric materials are currently being employed in microelectronic, microwave, and optoelectronic applications. For example, the large capacitance per unit area requirement in high density DRAMs is accomplished using high ε_r ferroelectric materials such as $Ba_{(1-x)}Sr_xTiO_3$ (BSTO) [17]. In the past few years, the use of high-permittivity ferroelectric materials in microwave devices has been widely investigated due to the increasing need for higher power, smaller size, lighter weight, and lower cost, frequency-agile components [18–21]. Examples of applications in the area of microwave engineering include field-dependent varactors, tunable resonators, phase shifters, frequency-agile filters, variable-power dividers, and variable-frequency oscillators [18–21]. Ferroelectric materials are also being investigated for nonlinear applications such as harmonic generators, pulse shaping, mixing, and parametric amplification [22]. Electronically tunable filters can be constructed and applied to interference suppression, secure communications, dynamic channel allocation, signal jamming, and satellite and ground-based communications switching. The use of high-permittivity ferroelectric ceramics and films requires the materials possess frequency stable permittivities and low dielectric losses. For microwave applications, $\tan\delta$ is lower in the paraelectric state ($T > T_c$) due to the absence of hysteresis loop. Operating just above T_c can yield large tunabilities.

BSTO is being investigated for applications in tunable microwave devices and circuits [17–20]. BSTO is a continuous, solid-solution ferroelectric with compositions between the perovskites $BaTiO_3$ and $SrTiO_3$. For instance, the Curie temperature (T_c) of BSTO decreases almost linearly with increasing Sr concentration while the index of refraction is changed by varying the solid solution composition and the deposition rate. Their nonlinear dielectric and optical properties can be tailored over a broad range to accommodate various applications. The BSTO (60/40) thin-film ferroelectric, with a T_c of 290 K, is a promising ferroelectric for room temperature tunable devices.

1.2.3. Photonics

Ferroelectric thin films are attractive for tunable photonic circuits due to the nonlinear electric field dependence of their refractive index at optical frequencies. Bias dependent structural changes in the ferroelectric thin films (ferroelasticity) are also attractive for tunable waveguides, modulators, and second harmonic generation [25]. Electric field dependence of birefringence and ferroelasticity makes them attractive for integrated optics applications. The thin-film form of ferroelectrics reduces the voltages necessary for achieving the required electric fields, as compared to the use of bulk ferroelectrics.

1.3. Nanophase Ferroelectrics

1.3.1. Background of Nanostructures

Nanostructured materials have dimensions that are measured on the length scale of a nanometer (one billionth of a meter) and have special properties due to quantum size and surface effects [26]. The motivation behind nanostructure engineering is the development of materials with novel properties through the controlled synthesis and assembly of the material on the nanoscale. No matter what synthesis approach is used there is a common premise for the fabrication of nanostructured materials. Fabrication of nanostructured materials allows the control and potential enhancement of the properties of the material, for example, ε_r, dielectric loss ($\tan\delta$), and Curie temperature, T_c. Nanostructure engineering allows miniaturization of devices to a critical size, below which the material loses its desired characteristics. There are a variety of approaches that can be used in the synthesis of nanostructures (Fig. 4).

A bottom-up approach uses advanced chemical techniques to mix precursor materials at the molecular level. Processing then allows growth of the material for the desired nanostructure; an example is sol–gel [27]. In the top-down approach, for example pulsed laser deposition (PLD), the

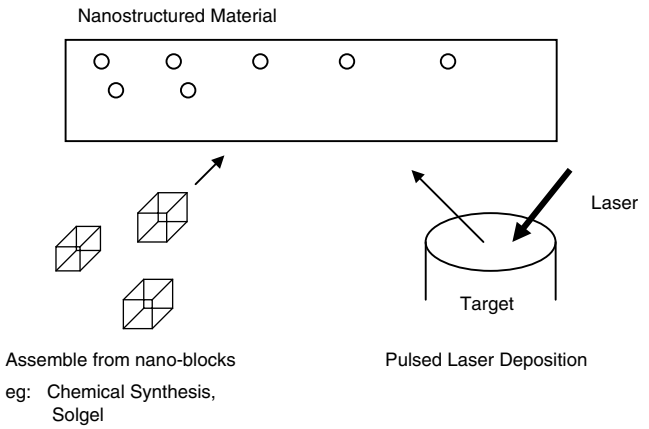


Figure 4. Synthesis of nanostructured material, the bottom-up approach, the top-down approach.

nanostructured material is deposited by the ablation of material from a target with a short laser pulse. The nanostructure of the deposited film can be controlled by varying the following PLD deposition parameters: chamber ambient gas, laser beam energy density, substrate temperature, deposition time, target system parameters, pulse repetition rate, pulse length, substrate-to-target distance, laser beam wavelength, target–substrate relative geometric arrangement, and postdeposition annealing [28–32].

Examples of materials with unusual properties at nanoscale include Si nanocrystals. Research in this area has gained momentum because of the potential of designing Si-based nanostructured materials that possess novel electronic and optical properties for use in new devices for optical communication. Quantum confinement in nanophase Si has revealed new physics for use in novel applications. Si is an indirect bandgap material and quantum confinement in this type of material can enhance luminescence, which is forbidden in bulk Si. Coalesced Si nanocrystals can be prepared by the laser vaporization-controlled condensation (LVCC) method. The Si fabricated nanocrystals have shown red photoluminescence upon excitation with visible and ultraviolet (UV) light. Since the unprecedented discovery of photoluminescence (PL) from porous Si, interest has grown in the synthesis and electrical/optical properties of Si nanocrystals because they are believed to be the luminescent centers in porous Si [33].

Very intense and highly reproducible PL spectra for SiO_x films have also been obtained by laser ablation of Si targets [34]. Patrone et al. were able to show that Si nanoclusters produced by laser ablation exhibited PL properties within a spectral region extending from the near UV to the near infrared (IR). The PL spectra were dependent on nanocluster size [35, 36]. The development of nanostructured materials has the potential to change present application of materials.

1.3.2. Size Effects in Ferroelectrics

For microwave applications, nanostructured ferroelectric films need to have a low dielectric constant ($25 < \varepsilon_r < 500$), a low loss tangent (or dissipation factor) ($0.001 < \tan\delta < 0.005$), and a high dielectric tunability [$((\varepsilon_{r0} - \varepsilon_{rV})/\varepsilon_{r0}) > 10\%$ where ε_{r0} is the ε_r at zero bias, ε_{rV} is the ε_r at a bias voltage V]. One of the methods to reduce ε_r and the $\tan\delta$ and improve tunability in BSTO is to form composites of BSTO with other nonelectrically active ceramics such as MgO and $LaAlO_3$ (LAO) [37]. One mol% Mg doped BSTO (60/40) resulted in a ε_r of 420, $\tan\delta$ of 0.022 (MHz frequencies), and tunability of more than 20% in thin films deposited by the metal organic deposition technique on Pt substrates [37]. Five mol% Mg content in BSTO reduced the ε_r to 386 and loss tangent to 0.007 (MHz frequencies) and increased tunability to 17.2% [38].

The potential for tuning the microwave properties of the ferroelectric barium titanate has been demonstrated by McNeal et al. [39]. Through the control of grain/particle size and domain state, relaxation frequencies were shown to increase and loss tangents decreased with decreasing grain/particle size. A polycrystalline ceramic ferroelectric, $BaTiO_3$, having grain sizes of 14.4, 2.14, and 0.26 μm and powder–polymer matrix composites of sizes 1.33 μm, 0.19 μm, and ~66 nm, respectively, were used.

Zhang et al. studied finite-size effects in the ferroelectric solid solution $Ba_xSr_{1-x}TiO_3$. Their samples were characterized and analyzed by scanning electron microscopy (SEM), transmission electron microscopy, and X-ray diffraction. The Curie temperature decreased with a decrease of grain size. The Curie peak became broader and eventually disappeared. Three ferroelectric critical grain sizes of 317, 246, and 176 nm for $x = 0.3$, 0.5, and 0.7, respectively, were evaluated [40].

Jiang and Bursill have proposed a new phenomenological model to study size effects on the first-order phase transition of lead titanate ferroelectric particles. The proposed model would try to explain the experimental results that show the Curie temperature, tetragonality (c/a) ratio, latent heat, and soft-mode frequency all decrease with decreasing particle size for $PbTiO_3$, $BaTiO_3$, and $PbZrO_3$ [41].

Wang et al. report on the size effects on spherical ferroelectric particles described by the transverse Ising model taking into consideration the long-range interactions [42]. In this report the authors studied the size dependence of the mean polarization, the Curie temperature, the mean susceptibility of the particle, and the critical size of the particle. The Hamiltonian of the transverse Ising model has the form

$$H = -\Omega \sum (S_i)^x - \frac{1}{2} \sum J_{ij} (S_i)^z (S_j)^z - 2\mu E \sum (S_i)^2 \quad (1)$$

Ω is the transverse field, and $(S_i)^x$ and $(S_j)^z$ are components of a spin-$\frac{1}{2}$ operator at site i. J_{ij} is the exchange strength between sites i and j. μ is the effective dipole moment of site i and E is the external electric field. The authors summarize their results: with a decreasing particle size, the mean polarization and Curie temperature of the particle decrease to zero at a critical size of the particle. Here, there is size-driven phase transition from the paraelectric phase to the ferroelectric phase. Long-range interactions are dominant contributions to ferroelectricity.

1.3.3. Why Nanophase Ferroelectrics?

Ferroelectric materials are typically multidomained as a large depolarization field (due to buildup of surface charges) forces a single crystal to divide into several domains. At microwave and optical frequencies, these multidomained materials exhibit a large dielectric relaxation characterized by a decreasing ε_r and increasing $\tan\delta$. Mechanisms attributed to dielectric relaxation include piezoelectric resonance of grains and domains, domain wall vibration, and shear waves from domain walls contributing to high frequency dispersion of ε_r $(\varepsilon_r' - j\varepsilon_r'')$ and high $\tan\delta$ $(|\varepsilon_r''/\varepsilon_r'|)$ [43]. The dielectric relaxation, which is intrinsically dependent upon domain state, is in part dependent upon the microstructure. Previous studies in size effects have shown that: (i) Relaxation frequencies increased and $\tan\delta$ decreased with decreasing grain/particle size [43]. (ii) T_c decreases with decreasing grain/particle size [43]. (iii) The peak value of ε_r decreases [44]. (iv) Lattice constants change such that the unit cell becomes more symmetric [44]. (v) There exists a critical domain size below which the ferroelectric behavior vanishes [43]. (vi) Monodomain ferroelectrics can be obtained for a grain size below 200 nm [43]. (vii) Domain switching is quicker in smaller domains compared to larger domains but requires higher electric fields [43].

2. SYNTHESIS OF NANOSTRUCTURED BSTO

2.1. Sol–Gel

Sol–gel is a chemical process that permits molecular-level mixing and processing of starting materials and precursors in a relatively easy way and permits nanostructured films to be formed at low temperatures. The process is relatively inexpensive because it does not require a vacuum for deposition. A sol–gel deposited film is amorphous, which means there is a long diffusion process with sufficient time for diffusing atoms to move into and through grains from grain boundaries. This leads to a relatively uniform distribution of diffusant [45–50]. Nanocrystalline particles of $BaTiO_3$ < 30 nm have been successfully fabricated using the sol–gel method [51]. A flow diagram for the sol–gel fabrication of $(Ba,Sr)TiO_3$ films is given in Figure 5. Precursor solutions are prepared in a nitrogen dry box at room temperature. Barium diethoxide $[Ba(OC_2H_5)_2]$ and strontium diethoxide $[Sr(OC_2H_5)_2]$ are dissolved in dehydrated 2-methoxyethanol $(CH_3OCH_2CH_2OH)$ and stirred. Titanium tetra-I-propoxide $[Ti(OCH(CH_3)_2)_4]$ will be added to the barium and strontium solutions [52].

The reaction represents the synthesis for $t \geq 0$ °C:

$$Ba(OH)_2 * H_2O + Ti(OR)_4 \rightarrow BaTiO_3 \quad \text{(I)}$$

$$Sr(OH)_2 * H_2O + Ti(OR)_4 \rightarrow SrTiO_3 \quad \text{(II)}$$

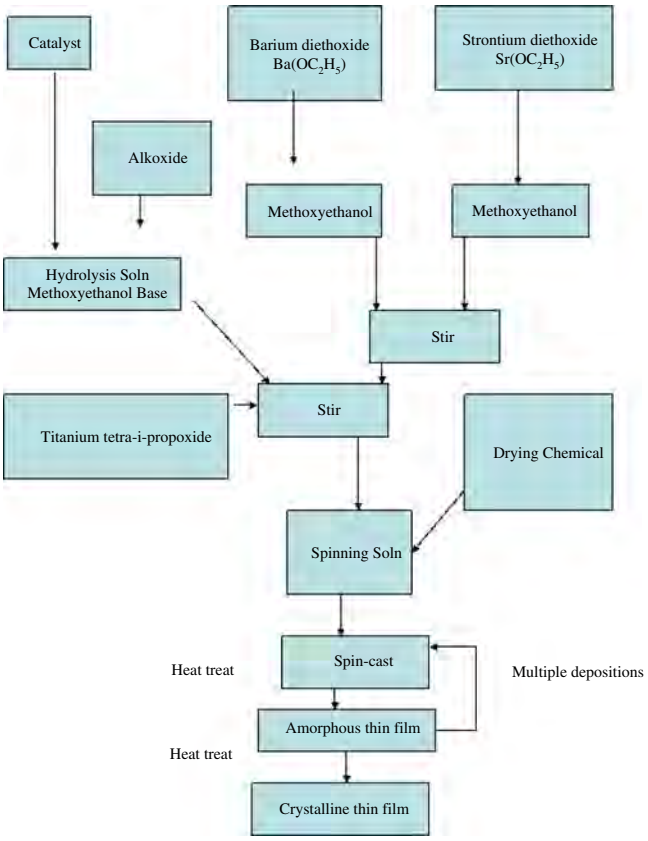


Figure 5. Flow diagram for the fabrication of sol–gel films.

2.2. PLD

PLD is recognized as a versatile and powerful technique for the fabrication of nanostructures. The principle of PLD utilizes an intense laser pulse passing through the optical window of a vacuum chamber. The laser beam is directed onto a solid or liquid target surface where it is partially absorbed. Above a threshold power density, target material is ablated and forms a highly energetic and directed luminous plume. The power density needed to produce the plume depends on the target material and its morphology as well as the laser pulse wavelength and duration. Film growth occurs at the substrate in an ambient gas environment where energetic adatoms from plume components can reproduce an epitaxial film with the same stoichiometry as the target [53, 54]. A reactive gas or ion source may be added to aid the growth process. The PLD plume can be a mixture of macroparticles, molecular species, atomic species, and ionic species. Pulsed laser deposition is popular in the deposition of oxide thin films because of the demonstrated deposition of high quality thin films and heterostructures [55]. The final quality of the deposited film depends on the complexity of the deposited material system and the ability to control approximately 17 process control variables according to a deposition recipe [56].

As in PLD of thin films as well as with PLD of nanostructures, the critical growth dynamics are controlled by the characteristics of the plume components' adatoms, the substrate/seed layer/film surface, and the ambient gas

environment. These characteristics can be summarized as follows:

1. Adatoms
   a. Mobility on film/substrate surface
      i. Kinetic energy
         1. Due to plume dynamics
         2. Due to heater temperature
      ii. Time before movement quenched by arrival of next pulse
   b. Density per laser pulse
   c. Differences between several plume component adatoms (generally they need to arrive in some order and density)
2. Substrate/seed layer/film
   a. Number and arrangement of nucleation sites on surface
   b. Lattice match of surface to desired stoichiometric growth structure
   c. Strength of component confinement inside substrate, seed/buffer and film layers (diffusion effects)
3. Ambient gas
   a. Required for incorporation into film structure (usually cannot be suitably added later)

For a more complete and detailed description of PLD, see Chrisey and Hubler [53].

2.2.1. Process Control

To adapt PLD as a manufacturing process for thin nanostructured films, the critical film-growth process variables must be identified and controlled. This permits *in-situ* control of the process which could enable intelligent processing or self-directed process control (SDC). SDC consists of making *in-situ* sensed observations of the process. Stark and Laube's [57] four steps for the development and implementation of SDC are

1. Identify appropriate sensors and the variables that control the deposition process dynamics.
2. Determine empirical relationships by experimentation, since no good analytical models exist at present.
3. Correlate test results with experimental results to develop qualitative relationships.
4. Incorporate qualitative relationships into the process baseline.

Many of the process variables in PLD, such as laser fluence and chamber ambient pressure are interdependent and are coupled to temperature-dependent variables like substrate temperature and plume velocity. Mathematically modeling the variables and their interdependence is very complex and has only partially been addressed [58–63]. Plume dynamics is one of the most important but least understood aspects of the PLD process. The plume is second only to the film growth dynamics in terms of complexity. SDC requires identification of the interrelationship of process variables involved. The 17 outer ovals in Figure 6, are at the environmental level and are considered process control variables. The three inner ovals are critical measurable parameters affected by several process-control variables. The dashed outlined process variable is a significant control variable not yet easily adjustable on excimer lasers. The octagons are subdivided regions of the deposition process where the control variables exert their effects.

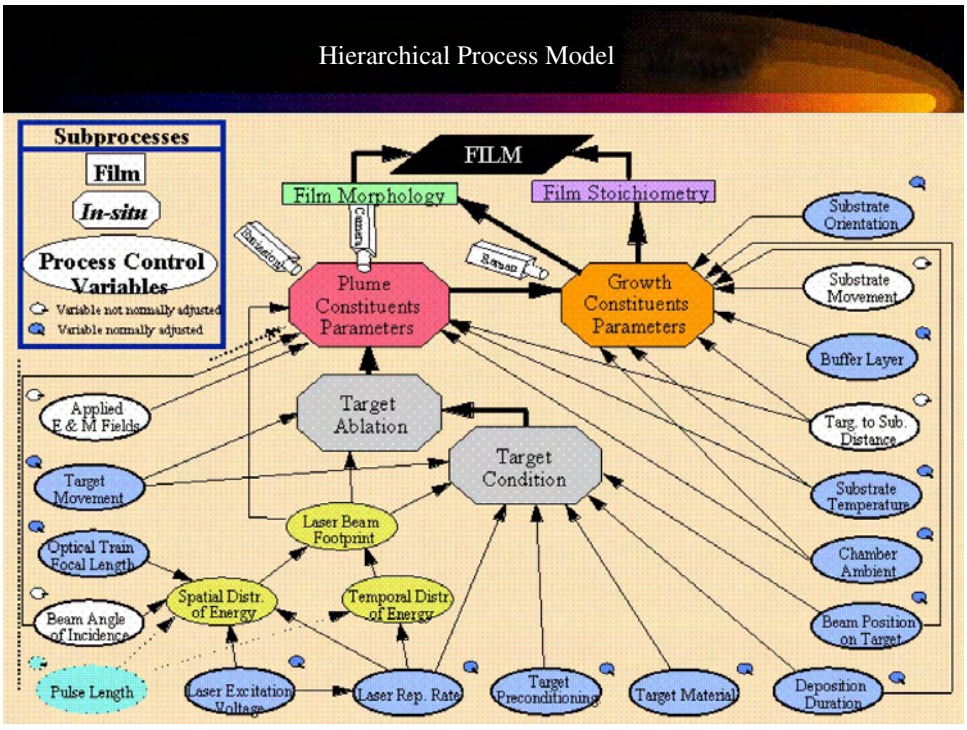


Figure 6. Hierarchial process model identifying process control variables and their complex interactions with a pulsed laser deposition system.

Appropriate sensors are positioned to monitor these critical regions. The knobs near the control variables indicate normally actuated/used when turned to a downward position. The emission sensor monitoring the plume can provide feedback activation of all but the top right three variables. These variables require a different sensor such as the Raman sensor depicted in Figure 6.

2.2.2. Versatility and Complexity of PLD

The versatility and complexity of PLD are evident by the number of process control variables that were noted in the hierarchical model. These deposition variables can be adjusted and each can strongly affect the basic components involved in the film/nanostructure growth depending on what part of deposition parameter space the deposition is in. Usually they are set according to a deposition recipe that is unique to the deposition system being utilized. Adjustment of these deposition variables is not recommended without some idea of the effect on the film growth. Biggers and Jones have developed a general list similar to that in the hierarchical model (to be published). These variables all affect the film growth through the adatoms, the substrate/seed layer/film, and the ambient gas at substrate/film and plume interface (where the film grows).

The following PLD process control variables strongly affect these adatoms, the substrate/seed layer/film, and the ambient gas environment during nanostructure growth.

1. The excitation voltage (kV) affects the target surface condition and film adatoms
   a. Controls total energy delivered to optical train. Some lasers may have a change in beam quality at a certain kV below which a current recipe will not produce quality films.

2. The optical train affects the target surface condition and film adatoms
   a. Controls the total energy (J) delivered to target
      i. Aperture reduces beam size (cm^2) and energy (J)
      ii. Affects beam quality
      iii. Each optical surface removes energy (J) from the beam
   b. Sets the beam footprint area (cm^2) on the target

3. Beam footprint (cm^2) affects the target surface condition and film adatoms
   a. At a set excitation voltage, controls power density (W/cm^2) applied to the target
   b. Affects the amount of target material removed per pulse
   c. Effects plume dynamics
      i. Expansion
      ii. Components' velocities

4. Substrate temperature (°C) above room temperature affects the film adatoms, ambient gas environment, growth surface, and diffusion effects
   a. Directly effects adatom mobility via heater surface temperature
   b. Effects plume dynamics by decreasing ambient gas density which lowers the number of component collisions with the ambient gas
      i. This decreases plume component cooling which increases arriving adatom kinetic energy (mobility)
      ii. Also increases plume expansion (reduces amount of adatoms arriving per area per pulse)
   c. Increases diffusion effects (impurity movement to grain boundaries, across interfacial surfaces, etc.)

5. Chamber ambient gas pressure (mTorr) and (type) affects the adatoms, substrate/film, and the ambient gas environment
   a. Affects the collision cross-section (probability of collisions)
   b. Affects the amount of energy transferred during collisions (cooling/slowing of plume components)
   c. May provide needed component to growing film
   d. Reduces or oxidizes the substrate surface

6. Target to substrate distance (cm) affects the adatoms and the ambient gas environment
   a. Affects kinetic energy of components (adatom mobility)
   b. Number of adatoms deposited per pulse (film growth dynamics)
   c. Plume envelope expansion (density of adatoms)

7. Deposition duration (min) affects target, film surface, and film
   a. Total amount of material delivered to substrate (thick films lose texture)
   b. Target surface dynamics (creation of cones, etc.)
   c. The diffusion range and amount of materials within the film/substrate system

8. Target system parameters include rotation rate (rpm), surface structure (polished or ablated), chemistry ($Ba_{0.5}Sr_{0.5}TiO_3$), and density (g/cm^3) which affect the film adatoms and film surface

9. Pulse repetition rate (Hz) affects film adatoms and film surface
   a. Time for adatom movement between pulses strongly affects film growth dynamics (longer times larger grains, shorter times probability of holes and outgrowth development, etc.)
   b. Amount of overlap between successive laser footprints (ablating from a solid or a liquid surface-effects kinetic energy of plume components, etc.)

10. Pulse length (ns) affects film adatoms
    a. Affects plume dynamics (longer pulses more plume initial ionization)
    b. The shorter the pulse length, the more ablative the plume generation

11. Wavelength (nm) affects film adatoms
    a. Affects the transfer of energy to target material (excimer may provide almost true ablation while IR causes target heating and material evaporation over a longer time frame)

b. Amount of energy available in beam footprint
c. Energy density and distribution available in the beam footprint

12. Postdeposition annealing affects film through diffusion effects and the ambient gas environment

2.2.3. Power and Flexibility with PLD

Grain size, surface smoothness, and film density are important characteristics for many thin film systems, especially superconducting and ferroelectric films. Many groups have been able to demonstrate some level of control over the final film characteristics to make the films more suitable for intended uses. They have utilized a subset of the deposition variables previously listed. Trananovic et al. utilized the pressure of the chamber ambient gas and the substrate temperature to control the surface morphology, grain dimensions, and grain density [64]. At low pressures (40 to 60 mTorr O_2), the YBCO *a*-axis grains were larger. The grain surfaces were flatter at 800 °C but rounded and better packed at 760 °C. Kim et al. also used ambient pressure to control film orientation and surface roughness [65]. They made a 30 nm template layer in 300 mTorr O_2 on Pt(111)/SiO_2/Si which produces an (110) orientation in the BTO film. They then deposited another 300 nm of BTO under 0.01, 100, and 400 mTorr. The orientation remained (110), the same as the seed layer, but the microstructure was quite different. The lowest pressure produced the densest and smoothest film, while the highest pressure produced the roughest surface and largest grains. Mukaida et al. used initial growth temperature to control YBCO grain size [66]. They found that grain size was logarithmically proportional to the heater temperature from 670 to 730 °C. In addition, if a desired grain size was set by an initial but short growth, then this grain size was retained if the temperature was changed for the remainder of the film growth. Biggers et al. found that laser repetition rate can strongly effect YBCO grain size and surface morphology while not diminishing other critical film characteristics such as critical transition temperature (T_c) and critical current density (J_c) [67]. The trend was that grain sizes decreased and the number of surface outgrowths increased as the repetition rate was incremented from 2 to 20 Hz. The number outgrowths continued to increase up to the limit of the laser at 40 Hz. However, the grain size slightly increased from that at 20 Hz. These are but a few examples of influencing film characteristics by manipulating the basic PLD deposition variables.

Emission Monitoring and Application of Process Control of PLD Emission spectroscopy (ES) permits the examination of plume sensitivity to almost any of the deposition process variables and is a critical element in achieving SDC. Pulsed laser deposition is very a versatile process especially for the deposition of nanophase thin films. Its potential is still constrained by the need to develop process control methodologies. Biggers et al. report the development of process control methodologies for a PLD system, Figure 7, [68] situated at the Materials Directorate, Air Force Research Laboratory, Wright–Patterson Air Force Base, Ohio. A PLD plume monitoring and control system was developed and utilizes a Lambda Physik 305i excimer laser (KrF, $\lambda = 248$ nm) with a 17 ns pulse, a variable-focus optical train, a beam scanning mirror, a Neocera vacuum chamber with multitarget holder, and substrate heater (see Fig. 7). Two optical emission sensor systems are attached to either side of the vacuum chamber to enable simultaneous monitoring of the same or different sections of the plume. The first maintains a fixed line-of-sight of a ~1 cm wide vertical slice of the passing plume manually set 38 mm from the target. The second system is similar but can be automatically positioned anywhere along the plume centerline from 11 mm away from the target all the way to the substrate heater. The two optical sensors in each system consist of photomultiplier tubes with inline narrow bandpass filters and slit arrangements to limit the viewed section of the plume. The filters are used to restrict monitored plume emissions to wavelengths associated with critical plume components (553 nm for exited Ba^* in $YBa_2Cu_3O_{7-\delta}$ depositions). The monitored emissions generate a time-of-flight (TOF) versus intensity signal as seen in Figure 8. The time difference between the start of the ablation process (the fireball) and the maximum intensity from the passing emitting plume component (main plume), the TOF setpoint, is sent to the computer.

This passed TOF provides the feedback for the computer to actuate from one to three deposition control variables (laser excitation voltage, chamber ambient pressure, and/or beam footprint size, process control variables 1, 5, and 3 respectively). Regulation of these deposition parameters once every two seconds to maintain the TOF setpoint provides the *in-situ* real-time process control. Also, linked through the computer is the control of the scanning mirror, laser repetition rate, and substrate temperature and recording of 27 deposition related variables, visible plume images, and control variable settings.

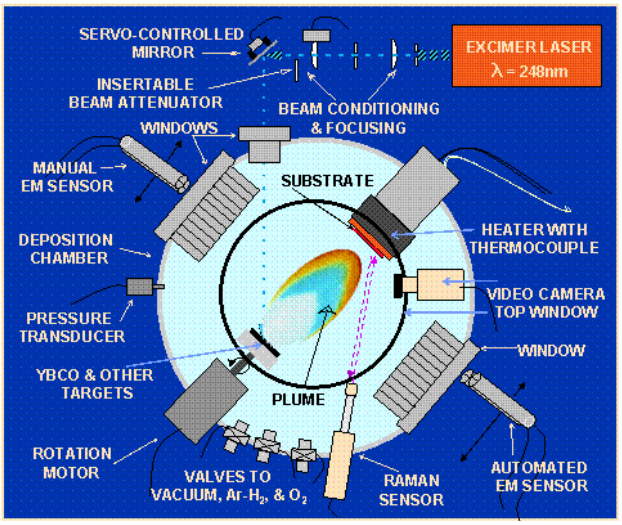


Figure 7. Automated *in-situ*, real-time, process-control pulsed-laser-deposition system. Real-time control based on feedback from emission sensors. Laser energy, chamber ambient, and/or beam footprint can be adjusted to maintain a TOF setpoint based on ES feedback.

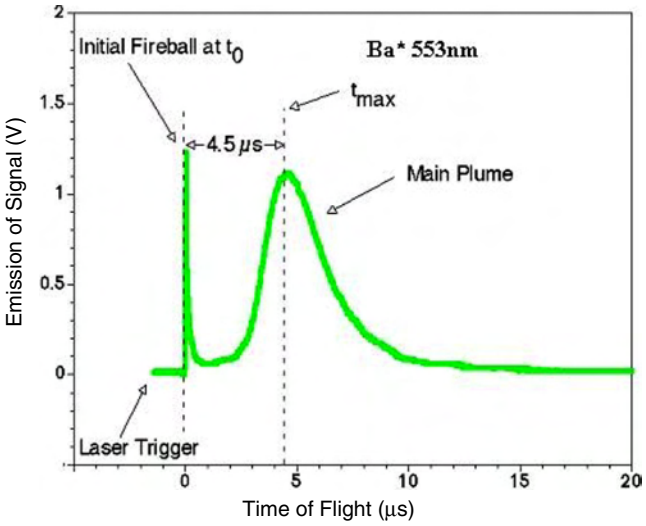


Figure 8. The emission signal from the Ba* component in the plume is generated when the material passed the monitored position. Note the sharp spike when the beam hit the target (fireball). This provides a timing mark for the start of the deposition process. The feedback TOF is the time between the fireball and the maximum intensity of the Ba* signal.

Emission Monitoring and Process Control for PLD of BSTO A number of groups have used ES to study plume dynamics [69–72] but only one group attempts *in-situ* real-time control based on ES feedback [68]. In the following figures, Biggers et al. show plumes from the deposition of BSTO on LAO substrates using the automated PLD system previously described and shown in Figure 9. A TOF of 4.5 μs, which in prior BSTO depositions [68] gave good results, was used as a starting TOF setpoint for the following depositions at different ambient O_2 pressures. TOF setpoints for various pressures from 19 to 300 mTorr were generated by attaining a TOF of 4.5 μs in 150 mTorr O_2 and then changing the pressure from 19 to 300 mTorr while recording the average TOF at each pressure, and holding a constant laser energy. Noting that 4.5 μs was recorded at 150 mTorr as the pressures were increased from 19 mTorr suggests that the desired energy density had been maintained. The TOFs and associated pressures were as follows: 3.60 μs @ 19 mTorr, 3.66 μs @ 38 mTorr, 3.95 μs @ 75 mTorr, 4.50 μs @ 150 mTorr, 4.95 μs @ 225 mTorr, and 5.45 μs @ 300 mTorr.

Figure 10 shows the variation of grain size with pressures ranging from 38 to 300 mTorr. The average grain diameters grow from ~36 nm at 38 mTorr to ~137 nm at 150 mTorr. Somewhere above 150 mTorr, the growth dynamics change and the grains begin to coalesce. Second depositions at the same deposition parameters produced similar results.

2.2.4. Nanoparticle/Nanocluster Formation within PLD Plume

The basic principles of nanoparticle formation using conventional pulsed laser ablation have been reported by Marine et al. [73]. (Haberland's description of clusters is given [74]: aggregates of atoms, or molecules, with size ranging from tens up to a few thousands of atoms.) Nanocluster synthesis using conventional PLD is usually performed at medium to high gas pressure, for example 0.2–10 Torr. Condensation starts within the plume during expansion and cooling. Within the plume condensed particles undergo multiple collisions with the ambient gas molecules. This leads to a stabilization of nanoclusters before arriving at a collection substrate surface. The size of nanoclusters can be varied by varying laser parameters of fluence, wavelength, pulse duration, ambient gas pressure, and the collection position from the plume centerline. Smaller clusters are scattered more than larger; hence more smaller can be collected off axis nearer the target. The authors [74] further define the process for nanocluster formation in four independent steps:

- *Cluster nucleation:* Nucleation is a function of the material's thermodynamic properties and the dynamic conditions of the plume and the plume components.

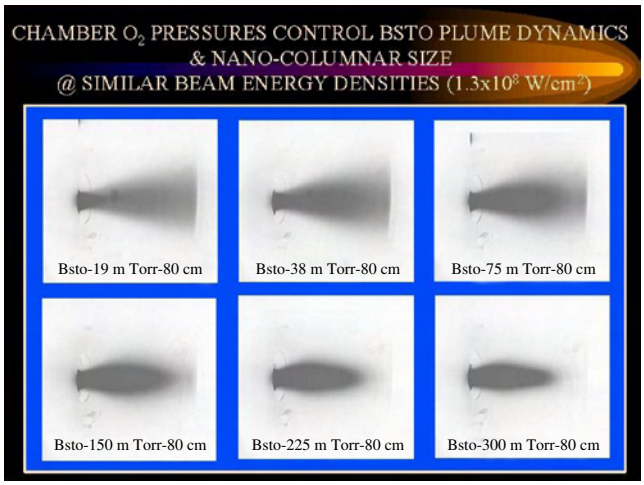


Figure 9. Visible wavelengths images of the PLD plume as the pressure was increased while holding a constant beam footprint energy density. Note the plume compacting and shortening, which is suggestive of more material being deposited per pulse and at lesser adatom energies.

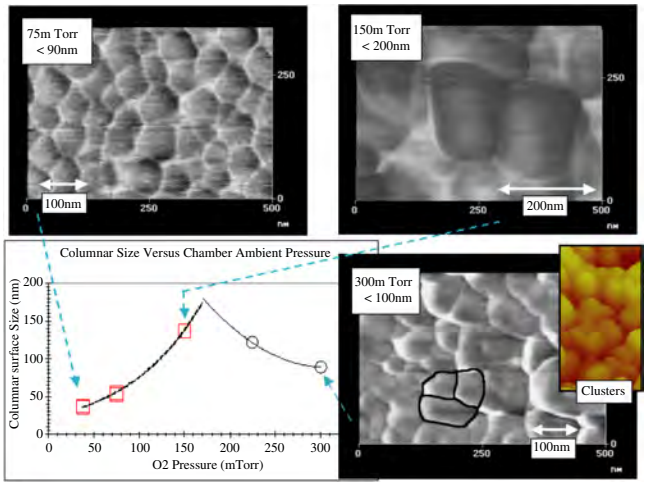


Figure 10. Frictional AFM images of BSTO surface morphologies as PLD ambient pressure increased. Note increasing grain size up to 150 mTorr of O_2, beyond which the grains begin to coalesce. The grain sizes fit a smooth curve up to 150 mTorr.

- *Growth:* Cluster growth occurs during the collisional expansion between plume particles (expansion dynamics vary with ambient gas effects). Expansion dynamics are influenced by ambient gas pressure.
- *Cluster cooling:* Because of the amount of energy generated during formation, clusters are stabilized by radiative, evaporative, and collisional cooling (collisional is the strongest with ambient gas at higher pressures).
- *Cluster deposition:* Deceleration of clusters must occur for cluster deposition or only fragments will be collected.

Deceleration is aided by growth in a background gas since this decreases significantly the kinetic energy of the ejected species by collisions with the gas species [74]. The mobility of the atoms arriving at the substrate surface is reduced relative to that in vacuum [74].

Nanocluster Size and Nanoparticle Formation It has been shown that the size of nanoclusters can be varied by the ambient gas pressure. Yoshida et al. [75] have shown that the diameter of Si particle size in the range 10–20 nm varied as a function of the pressure of the chamber ambient gas in the proportion $P^{1/2.8}$. Gonzalo et al. reported the influence of deposition parameters on nanocomposite materials produced by PLD [76]. Nanocomposite films of Cu nanocrystals embedded in amorphous Al_2O_3 were produced by PLD. The effect of laser fluence and ambient Ar gas was studied. The films were grown at room temperature using variable fluences and Ar pressures (10^{-6}–10^{-1} Torr). The authors suggested in the discussion of the experiment that nucleation and growth of the nanocrystals occurs on the substrate, not in the plume at 0.1 Torr when the diffusive region of the plume (the nucleation region) occurs at distances beyond the current substrate placement.

PLD Nanoparticle Condensation A phenomenological description of a condensation process is given by Pászti et al. for the formation of nanoparticles and nanocrystal networks prepared by PLD [77]. The authors postulate that in a high enough background pressure a shock wave is initiated by the collisions between target and gas atoms. This confines the plume and slows it down. In an inert gas atmosphere they describe four steps in nanoparticle formation:

- During the laser pulse, the plume components are emitted from the target.
- Atoms ablated from the target collide with ambient gas atoms primarily at the front of the expanding plume. Plume atoms thermalize, in 10–100 μs, after the laser pulse, while the shock wave develops.
- Nanoparticles form via homogeneous nucleation due to a supersaturation state, a result of the plume confinement and cooling behind the shock wave front.
- The first nanoparticles that are formed can coalesce incompletely which leads to the formation of aggregates in the plume and a nanocrystal network. This may coalesce completely leading to larger spherical particles. This continues until the number of primary particles become insignificant due to expansion.

This process explains the observed nanoparticle formation. There is a pressure range that accounts for thermalization in a limited volume of the plume. When decreasing the pressure, more plume component atoms avoid condensation and reach the substrate where heterogeneous nucleation begins. At the other end of the pressure range, increasing the ambient pressure increases the rate of thermalization decreasing the volume, where thermalization occurs. The level of supersaturation also increases. Thus the number and density of primarily condensed particles continues to increase. This leads to larger aggregates for incomplete coalescence and larger single particles in the complete coalescence. See [77] for a more detailed discussion.

2.3. Sputtering

Radio frequency (rf) magnetron sputtering is another popular technique in fabrication of complex dielectrics. The microstructure of the films depend on the sputtering pressure during the deposition, rf power, the target to substrate distance, and the substrate temperature. The scattering behavior of the sputtered components during deposition control the microstructure and crystal orientation of the growing film. Periaswamy et al. sputter deposited $Ba_{0.5}Sr_{0.5}TiO_3$ thin films on $Pt/TiO_2/SiO_2/Si$ substrates [78]. The films exhibited a capacitance tunability of 4:1. The higher tunability was attributed to (100) texturing of BST films and the biaxial tensile stress imposed by Si on BST making the polar axis oriented in plane [78]. Wang et al. demonstrated (111) oriented BST thin films (with 10% Sn doping in the Ti sites) on $Pt/TiO_2/SiO_2/Si$ substrates by rf magnetron sputtering [79]. The grain size increased with increasing deposition temperature, correlating with an increased dielectric constant [79].

2.4. Metal Organic Chemical Vapor Deposition

Metal organic chemical vapor deposition (MOCVD) is another powerful technique used in the semiconductor industry for large area growth of thin films. With suitable organic precursors, one can control the growth very precisely. The growth of ferroelectric thin films is dependent on the precursors and the substrate temperature. Gao et al. studied the effects of precursors and substrate materials on the microstructure and dielectric properties of BSTO thin films grown by MOCVD [80]. Two sets of precursors with quite different reactivities resulted in differing step coverage. Different reactivities apparently did not affect the microstructure, surface morphology, and dielectric properties of the stoichiometric BSTO thin films.

Tombak et al. studied BST thin films for variable capacitor (varactor) applications for rf and microwave frequencies [81]. BST films grown by the MOCVD method achieved a 71% tunability. The measured dielectric loss tangent was in the range of 0.003–0.009 at very high frequencies [81].

2.5. Molecular Beam Epitaxy

Molecular beam epitaxy is another popular process for preparing ferroelectric thin films. Ferroelectric $BaTiO_3$ (001) ultrathin film (10 nm) has been epitaxially grown on $SrTiO_3$ by laser molecular beam epitaxy [82]. The film was

observed by atomic force microscopy (AFM) and had a surface roughness of 0.2 nm.

Zhao et al. [83] report using laser molecular beam epitaxy to produce a series of $BaTiO_3$ thin films with thickness varying from 10 to 400 nm. The oxygen pressures were varied from 1.5×10^{-3} to 90 mTorr on $SrTiO_3$ (001) substrates. The films were highly *c*-axis or *a*-axis oriented with a surface root-mean-square (rms) roughness of 0.14 nm. The dependence of the crystallographic orientation on gas pressure was reported. Two mechanisms were reported as responsible for the effect: the residual hydrostatic compression from excess interstitials under low pressure and the residual hydrostatic tension resulting from excess vacancies under high pressure caused the crystallographic orientation.

3. SPECTROSCOPIC CHARACTERIZATION

3.1. X-Ray Diffraction

X-ray analysis reveals lattice structures as well as crystal structure. When X-rays are directed at a crystalline material they are diffracted by the planes of atoms or ions within the crystal. The diffraction angle depends on the wavelength of the X-rays and the distance between adjacent planes.

X-ray diffraction (XRD) analysis has been used to study tetragonality and lattice parameters of ceramic specimens of $BaTiO_3$ [84]. The ceramic materials used were prepared from commercial powders. Coarse grain $BaTiO_3$ (CGBT) polycrystalline ceramics of size 1–20 μm were processed by conventional solid state sintering. Small grain ($\sim$1 μm) ceramics of $BaTiO_3$ (SGBT) and fine grain ($\sim$0.02 μm) ceramics of $BaTiO_3$ were prepared using hydrothermal starting powders. The powders were dry pressed into pellets and further processed using cold isostatic pressing, then hot pressed, and annealed. Grain size was determined by SEM. X-ray diffraction analysis revealed a decrease in tetragonality as the grain size decreased from coarse to fine. The decrease in tetragonality was attributed to internal stress in the crystal [84]. The large grain size would allow for compensation for homogenous stress. The small grain ceramic allowed for partial internal stress relief. It was proposed that the fine grain size was probably too small to allow domain twinning to occur keeping high internal stresses throughout the grain which would reduce tetragonal distortion [84]. (Mechanical stress can be compensated for in a material through *twinning*. The stress can produce atomic displacements such that on one side of a plane, the twin boundary, atoms are located in mirror image positions of atoms on the other side of the boundary.) Diffraction peaks were obtained from $2\theta = 20°$ to 80°. The (200) reflections were observed for the three ceramic samples. The CGBT sample showed splitting of the reflection. The SGBT sample showed some splitting of the (200) reflection.

Process parameters for PLD play an important role in the microstructure of the thin film. The investigation by Kim et al. [85] reported tailoring the preferred orientation and microstructure of $BaTiO_3$ thin films deposited by PLD. The microstructure was affected by adjusting oxygen pressure during deposition. A KrF excimer laser was used with a pulse rate of 5 Hz for film deposition. A power density of 2 J/cm^2 was used. The target to substrate distance was 42 mm. The barium titanate pellets were prepared by isostatic pressing. Films were deposited on $Pt/SiO_2/Si$. Oxygen was introduced during the film deposition. Crystallographic structures and microstructures of the films were studied using XRD and SEM. The XRD patterns of $BaTiO_3$ showed that films deposited at 1 mTorr were oriented to (100). At oxygen pressure lower than 1 mTorr, dense and smooth films were obtained. Film roughness increased with increasing oxygen pressure.

Vijayalakshmi et al. [86] reported the growth properties of Si films deposited on Si substrates using PLD. A KrF excimer laser beam was focused to a 100 μm spot diameter on a Si film. Because of the directional nature of the PLD plume, the substrate samples contained regions that varied in thickness. There were three main regions, region I with estimated thickness 200 nm, region II with estimated thickness 400 nm, and region III with estimated thickness 800 nm. The three regions contained micrometer sized islands. Nanometer size clusters were scattered between larger islands. The authors report on a 2θ scan of the Si film; the XRD spectra revealed a sharp peak at 69.3° which indicates a higher order of (100) Si crystal. Approximately 83% of the crystallinity of the nanoclusters was seen to be from (111) and (220). The three regions I, II, and III of the deposited films showed the same 2θ peaks. The intensity of the XRD spectra varied. The average size of clusters increased going from region I to region III and in region II the average cluster size was 25 nm. The XRD peak intensity decreased going from region II to region III and the width of the peaks remained the same.

Pulsed laser deposition, using the system described in Section 2.2, was used to deposit 300 nm thick films of $Ba_{0.6}SrO_{0.4}TiO_3$ on substrate $LaAlO_3$. The microstructure of the film was varied by changing the ambient gas pressure while holding the energy density constant. At 38 mTorr, the particles were distinct and had average diameter 37 nm. As the pressure increased in steps from 38 to 150 mTorr the particle's grains remained distinct. However, by 225 mTorr, the nanograins began to coalese into clusterlike structures. The XRD spectra show that as the clusters form, the intensity of the (200) reflection for BSTO and the intensity for the (200) reflection of LAO reverse in relative magnitudes. The (200) reflection for BSTO is now greater in magnitude than the (200) reflection for LAO perhaps indicating the clusters form a better surface coverage of the substrate. Figures 11, 12, and 13 show the change.

3.2. Scanning Probe Microscopy

Scanning probe microscopy (SPM) is a class of surface characterization techniques utilized to study the physical, chemical, or electrical properties of surfaces and interfaces on the nanometer-to-micrometer scale [87, 88]. SPM technology provides measurement of three-dimensional surface topography, atomic resolution imaging of conductive and nonconductive materials, probing of various types of surface forces (friction, magnetic, electric) [89], unique analyses of surfaces (electrochemical reactions, surface elasticity/hardness, surface adhesion, surface charge density, surface energy, viscoelastic properties) [89], and surface

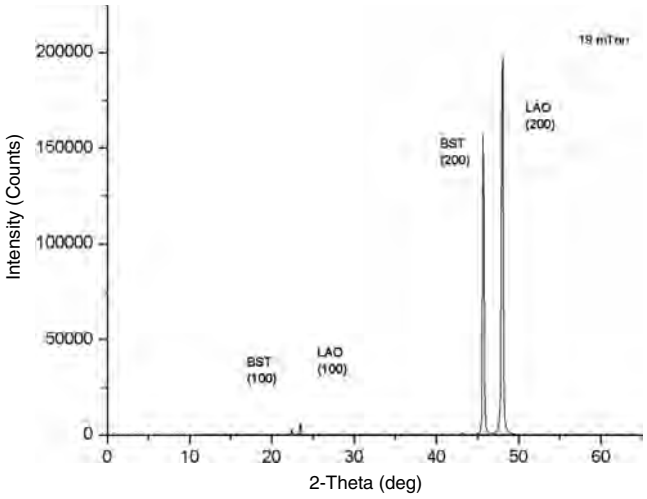


Figure 11. X-ray diffraction pattern for a BSTO film deposited on LAO at 19 mTorr oxygen pressure, radiation = 2θ(Cu $K\alpha$).

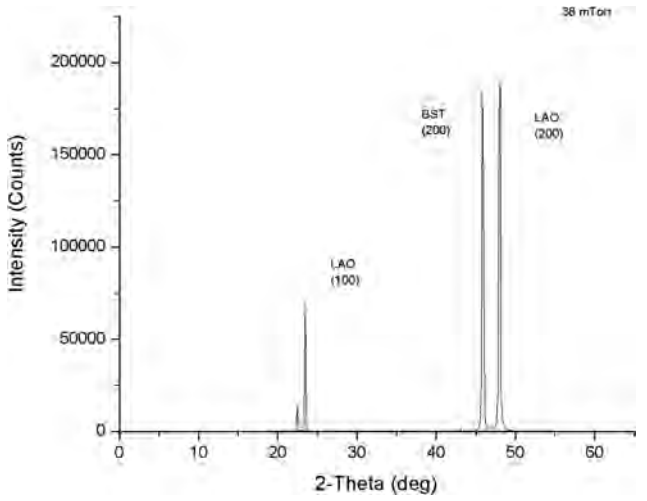


Figure 12. X-ray diffraction pattern for a BSTO film deposited on LAO at 38 mTorr oxygen pressure, radiation = 2θ(Cu $K\alpha$).

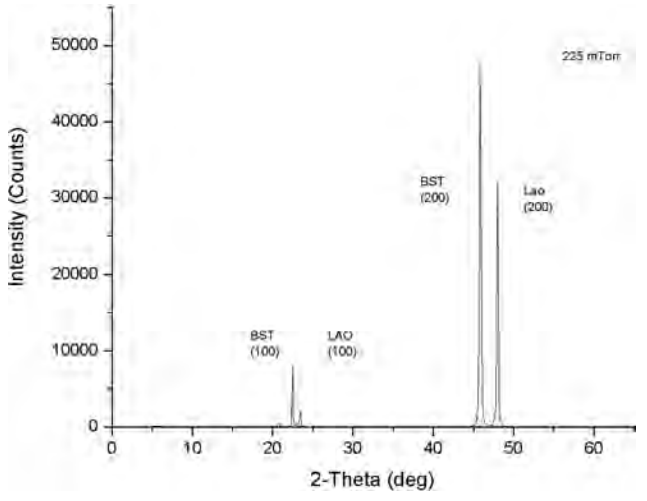


Figure 13. X-ray diffraction pattern for a BSTO film deposited on LAO at 225 mTorr oxygen pressure, radiation = 2θ(Cu $K\alpha$).

modification and manipulation [87]. The various modes of operation (scanning tunneling, contact, noncontact, tapping mode) enable the versatile nondestructive analysis of a wide variety of materials surfaces including ceramics, metals, polymers, biomolecules, semiconductors, superconductors, ferroelectrics, and composites. This makes SPM an essential tool in the characterization of nanostructures.

In general, SPM techniques involve scanning a sharp tip within a very close distance (in the near field) across a material surface, allowing the probing of specific physical, chemical, or electrical interactions that occur between the tip and the material surface [90]. The resolution of a given SPM technique depends on the size, shape, and sharpness of the tip. A smaller or sharper tip probes a smaller surface area and results in better lateral resolution [90]. Scanning probe microscopes measure and image surfaces in all three dimensions (x, y, z), with x and y topographic resolutions of 2–10 nm for AFM, 0.1 nm for scanning tunneling microscopy (STM), and a z resolution of 0.1 nm [88]. Scanning tunneling microscopy, the first SPM technique, utilizes a very sharp conductive tip to produce real space images of atomic arrangements on highly conductive flat surfaces. Typical materials for STM probes are gold, tungsten, and platinum/iridium. The STM tip is mounted on a system of piezoelectric drives, which provide movement in three dimensions [91]. In STM measurements the very sharp tip is brought within 0.1 nm of the conductive material surface while a voltage (bias) in the range of 0.1–1 V is applied between the tip and the sample. This close distance causes the outer electron orbitals of the tip and sample to overlap. The voltage causes the electrons to tunnel from the tip into the unoccupied energetic states of the sample [90, 91] producing a tunneling current, which is used to map the topography of the sample surface. The STM can be operated at either constant height mode or constant current mode. In the constant height mode the tip is scanned over the sample surface at a fixed height while the current is measured. A topographical image is obtained by mapping the current tip–sample separation relationship. In the constant current mode, the tip-to-sample separation is varied using a feedback loop to maintain a constant current as the tip scans across the sample surface resulting in a image based on the position of the tip.

The STM provides subangstrom resolution but it is limited to imaging highly conductive materials. Because of this limitation AFM was invented to map the three-dimensional topography of nonconductive samples. AFM utilizes a very sharp tip (silicon nitride or silicon) placed on the end of a cantilever to probe and map the topography of both nonconductive and conductive surfaces. The cantilever has a low spring constant on the order of 1 N/m. The most common modes of AFM are contact mode AFM and tapping mode AFM. In contact mode AFM, as the tip is scanned across a sample surface the force between the tip and the sample (10^{-9} N) is measured by monitoring the deflection of the cantilever, which occurs as a result of the tip pushing against the sample surface [87, 88]. A pieozoelectric tube scanner is utilized to either scan the tip in a raster pattern with respect to the sample or to scan the sample with respect to the tip [92]. As the scanning progresses, interactions between the tip and the sample are monitored by

reflecting a laser beam off the back of the cantilever into a four-part photodiode detector. Changes in the cantilever deflection are monitored by detecting the difference in the photodetector output voltages that result as the position of the laser beam on the photodetector changes [93]. The laser beam movement is monitored both vertically and horizontally. As the tip scans across the sample and encounters surface features, the upward and downward movements of the tip shift the reflected laser beam between the lower and upper portions of the photodetector and create voltage differences which are electronically converted into height information. A feedback loop maintains a constant cantilever deflection and thus a constant force (0.1–100 N) during scanning by moving the piezoelectric scanner vertically at each x, y datapoint, to maintain a constant voltage difference between the upper and lower portions of the photodetector [92]. The distance the scanner is moved vertically at each x, y datapoint is utilized to render a topographical image of the sample surface. Lateral movements of the beam between the left and right portions of the photodetector create voltage differences, which are converted into frictional information. In tapping mode AFM the cantilever (with a silicon tip) is oscillated at its resonant frequency and lightly taps the surface of the sample during scanning. A feedback loop maintains a constant rms oscillation amplitude by moving the piezoelectric scanner vertically at each x, y datapoint [92]. The vertical movement at each datapoint is used to render a topographical image. Tapping mode AFM is preferred over contact mode AFM when imaging soft fragile samples like polymers, biomolecules, and adhesives because of the absence of lateral shear forces encountered with the contact mode that often damage the sample. Tapping mode AFM is also an excellent tool in the imaging of other classes of materials such as silicon wafer surfaces, superconductors, metals, and insulators.

Lowndes et al. [94] report the characterization of Si and ZnTe nanoparticles produced by PLD using AFM. The size distribution measurements were made for nanostructures too small for an accurate high-resolution (field emission) scanning electron microscope (HRSEM, Philips XL30/FEG) in secondary electron mode. The heights (approximating the diameters) for nanostructures too small for accurate HRSEM measurements, all of the Si and some ZnTe, were made using the tapping mode (TM) AFM. A Nanoscope III (Digital Instruments) was used to image 500 nm^2 areas in two different regions of each substrate. The TM-AFM had vertical resolution of ~0.05 nm. The horizontal resolution in the study was limited by the size of the imaging tip. The nanoparticle size distribution was obtained by statistical analysis of the TM-AFM images [89]. The size distributions of Si and AnTe nanoparticles, produced by low energy density ArF (193 nm) pulsed laser deposition in He gas, were correlated to the gas pressure and target–substrate separation. Their AFM measurements of the Si nanoparticle size distribution showed that for He pressures in the range 1.5 to 10 Torr, larger nanoparticles were more numerous closer to the target. With increasing substrate to target distance, smaller nanoparticles formed. The authors point out that the mean Si nanoparticle size was maximized for $P \approx 6$ Torr while in contrast Yoshida et al. [95] and Makimura et al. [96] found the mean diameter of Si nanoparticles increased monotonically with the He pressure in the same pressure range. When correlated with their visual observations of the plume, the AFM data showed that for any He pressure, the largest nanocrystals were deposited beyond but close to the end of the luminescent plume. Nanocrystal size distribution shifted toward smaller average diameters for larger substrate-to-target distance. The authors note this trend had not been identified in earlier studies.

Contact mode atomic force microscope images (Fig. 18) of BSTO thin films deposited onto LaO substrates with the PLD system described in Section 4.2.4 were obtained using a Digital Instruments MultiMode SPM with a NanoScope IIIa controller in air at ambient temperatures at Wright–Patterson AFB in the Materials Directorate. Scan sizes for the images range from 500×500 nm to 167×167 μm. Scan speeds for the images were between 0.5 and 3 Hz. Images were obtained from multiple locations on each BSTO thin film sample. All analyses of the nanostructure dimensions and surface roughness measurements were performed on flattened images. Multiple 500×500 nm areas from several 1×1 μm images were used to determine nanostructure widths and heights for each thin film. Surface roughness measurements for each thin film were obtained from multiple 1×1 μm areas within several 5×5 μm images. The images in Figure 14 show the particle size increasing with increasing O_2 ambient pressure from 38 to 150 mTorr.

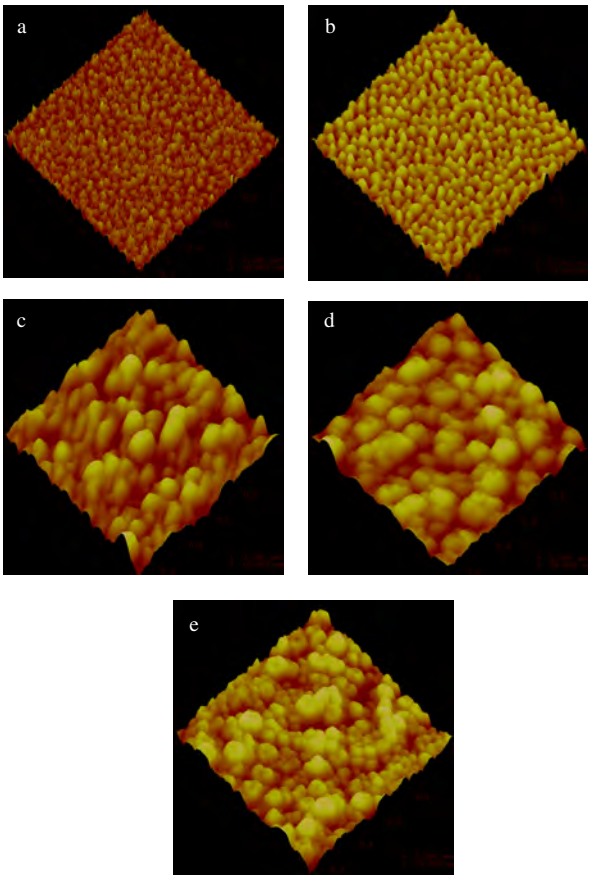


Figure 14. AFM images for (a) 38, (b) 75, (c) 150, (d) 225, and (e) 300 mTorr.

Above 150 mTorr, clustering or coalescing of the grains was observed starting at 225 mTorr.

4. PROPERTIES

4.1. Electrical Properties

4.1.1. Effect of MgO-Mixed BSTO Bulk Ceramics and Thin Films on Dielectric Properties for Tunable Microwave Applications

Chang and Sengupta investigated the dielectric properties of $Ba_{0.6}Sr_{0.4}TiO_3$ and 1.0, 20.0, 40.0, and 60.0 wt% MgO mixed with bulk ceramic and thin film BSTO [97]. The BSTO and MgO-mixed BSTO bulk ceramics were prepared by a standard solid-state reaction process described by Sengupta et al. [98]. BSTO and BSTO–MgO composite thin films (~0.03 μm) were grown on (100) MgO single-crystal substrates by PLD. The process is described by Grabowski et al. [99]. Chang and Sengupta observed that the grain size of BST bulk ceramic was relatively large (~10–20 μm) compared to 1.0 wt% MgO-mixed BSTO bulk ceramic (~3–5 μm) and added that MgO is an observed inhibitor of grain growth in titanate ceramics. The SEM images of 1.0, 20.0, 40.0, and 60.0 wt% showed no significant change in grain size. The thin films showed a similar trend as in the grain size for bulk ceramics with change in wt% of MgO. The grain size of the thin film was much smaller than the grain size of the bulk ceramics. The grain size for pure BST thin film was ~100 nm and for MgO-mixed BST was ~50 nm (seen in the SEM images).

The dielectric constant of the bulk ceramics and thin films drops at 1.0 wt% MgO-mixed BST and gradually decreases with an increase of wt% MgO. The dielectric tuning of bulk ceramics and thin films shows a similar trend with respect to the MgO mixing content. The dielectric constant and dielectric tuning for thin film remains less than bulk constants except at ~50 wt% where there is a reversal in magnitude of bulk and thin film dielectric constant. The dielectric Q of the thin films deposited using pure BST and MgO-mixed BST bulk ceramics was given as a function of applied dc electric field:

- Thin films 20.0, 4.0, and 60.0 wt% MgO-mixed bulk ceramics dielectric Q:

$$Q \sim 75 \text{ at } 0 \text{ V}/\mu\text{m to } 150 \text{ at } 16 \text{ V}/\mu\text{m}$$

- Pure BST and Mg-substituted BST thin films:

$$Q \sim 20 \text{ at } 0 \text{ V}/\mu\text{m to } 75 \text{ at } 16 \text{ V}/\mu\text{m}$$

In summary, the authors have shown that Mg substitution into BST modifies the microstructure of bulk ceramics and thin films significantly. There is a shift in the cubic-tegragonal phase transition peak (T_C) to a lower temperature. This causes an abrupt decrease in the dielectric constant and tuning.

4.2. Physical Properties

4.2.1. Effect of Porosity on Permittivity and the Effect of Temperature on Permittivity and Tunability

The influence of the microstructure (particularly porosity and grain sizes) of $Ba_{0.6}Sr_{0.4}TiO_3$ on the dielectric properties of bulk ceramics and thick films was studied by Zimmermann et al. [100]. Their paper investigates BSTO bulk ceramics and thick films with different porosities and grain sizes. Dense and porous BSTO bulk ceramics were prepared by the mixed oxide method [100]. The authors mention that it has been shown that the permittivity of barium titanate has been reduced by increasing the porosity of admixing different materials [101, 102]. The temperature dependence of permittivity has been shown to decrease by reducing grain size [103, 104].

The temperature dependence of the permittivity and loss tangent of dense and porous BSTO bulk ceramics was shown. The permittivity curve of the dense bulk ceramic showed three phase transitions from cubic to tetragonal (275 K), tetragonal to orthorhombic (220 K), and orthorhombic to rhombohedral (175 K). The permittivity curve of porous bulk ceramic showed one broad peak and reduced permittivity.

In comparing the dielectric properties of the porous ceramic with the thick film, the thick film showed significantly smaller permittivity which decreases by increasing porosity. At room temperature the porous ceramic showed a tunability of 8%. The dense bulk ceramic had ~16% tunability. The thick films were almost temperature-independent and had a tunability of 70–80% at a maximum temperature and 45–70% at room temperature.

4.2.2. Dielectric Properties of Sol–Gel-Derived $MgO:Ba_{0.5}Sr_{0.5}TiO_3$ Thin Films

Jain et al. investigated the dielectric constant and the dielectric tunability sol–gel-derived $MgO:Ba_{0.5}Sr_{0.5}TiO_3$ thin-film composites [105]. In the study, MgO:BSTO films were fabricated by sol–gel to synthesize heterostructured thin films. The X-ray diffractograms showed that both pure and heterostructured films were primarily (100) oriented and crystallized into a perovskite phase. From cross-sectional SEM micrographs, the heterostructured films showed better density and low porosity compared to pure BST thin films. The dielectric constant and the dielectric loss tangent of the heterostructured films were lower than pure BST thin films especially in the ferroelectric state [105]. The authors mentioned that the significant reduction of loss tangent in the heterostructured BST thin films is attributed to the MgO solubility in BST and microstructural modifications induced by the MgO layers [105].

4.3. Microwave Properties

Determining the dielectric properties of the ferroelectric thin films at microwave frequencies is critical for developing tunable components such as phase shifters and filters. To determine the relative dielectric constant (ε_r) and the loss tangent ($\tan\delta$) of the ferroelectric thin films, coplanar

waveguide (CPW) transmission lines and resonators (loaded as well as unloaded) have been designed for on-wafer probing. With measurements of scattering parameters on the CPW lines and resonators in conjunction with theoretical quasi-static conformal mapping analysis, one can determine the dielectric properties of the ferroelectric thin films, including the frequency dependent $\tan\delta$. The scattering parameters (S-parameters) are useful for microwave frequencies as they are based on a traveling waves concept. The S-parameters can be determined based on appropriate impedance matching at both the source and load ends. The S-parameter S_{21} is the ratio of the output power to the input power in dB and is also called the insertion loss of the circuit under measurement. S_{11} is the ratio of the reflected power to the input power in dB. The S-parameters are complex quantities with a magnitude and phase. When we measure S-parameters for a CPW line, the magnitude of S_{21} is an indication of the loss in a line, and the phase of S_{21} is an indication of the electrical length of the line. For more information on S-parameters, please refer to a modern textbook on microwave engineering.

Once the ferroelectric thin films are deposited on the substrates, test structures are either photolithographically defined or samples are coated with a silver or gold conductor by sputtering, thermal evaporation, or PLD through the shadow mask. The minimum thickness of the metal layer required is approximately three times the skin depth of the metal at the lowest frequency of measurement. The CPW lines and resonators are very easily probeable using ground–conductor–ground probes.

To understand the effect of inserting the ferroelectric thin film in the CPW structure, tunable components were modeled such as a CPW transmission line and CPW linear resonators. The cross-section of the modified CPW structure, consisting of a conductor/ferroelectric/dielectric configuration, is shown in Figure 15. The structure consists of a dielectric substrate (typically LAO of 254 μm thick), a BSTO ferroelectric thin film layer (thickness t typically 300 nm), and a 2 μm thick Au/Ag thin film for the conductor (centerline) and ground lines (adjacent parallel lines).

The important geometrical dimensions are indicated in the Figure 15. The critical design parameters for the tunable circuits are the characteristic impedance (Z_0) and the effective dielectric constant (ε_{eff}) which are both a function of the electric field dependence $\varepsilon_{r\text{FE}}$ of the ferroelectric thin film. Important geometrical parameters controlling these parameters are the width of the center conductor (W), spacing between the ground lines (S), thickness of the ferroelectric thin film (t), and the thickness of the substrate (h). The characteristic impedance of the line is inversely proportional to the ratio of W/S. The dielectric properties of the ferroelectric thin film and the thickness of the ferroelectric film are expected to determine the overall insertion loss of the circuit. The transmission lines were designed for characteristic impedance of 50 Ω in the absence of the ferroelectric layer.

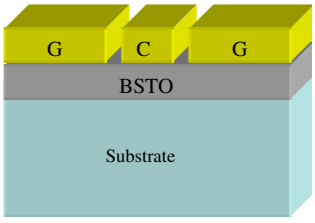


Figure 15. Cross-section of the modified ferroelectric tunable CPW configuration. The G–C–G refers to ground–conductor–ground.

The methodology in determining ε_r and $\tan\delta$ for the ferroelectric thin films at microwave frequencies is (i) Determine the effective dielectric constant of the CPW structure using the phase of the S_{21}. $\varepsilon_{\text{eff}} = [(\text{total phase of } S_{21} \text{ in degrees}) * \lambda_0/(360 * l)]^2$ where λ_0 is the wavelength in free space, and l is the length of the CPW line. (ii) Use a conformal mapping technique based on quasi-static analysis to determine the ε_r of the ferroelectric thin film [23]. Another method is to use simulation tools such as Sonnet and plot the effective dielectric constant versus the relative dielectric constant of the ferroelectric thin film, for the frequency of interest and for the specific film thickness [25]. Once we obtain ε_{eff}, we can extract the value of ε_r from such a plot. (iii) By measuring the quality factor of the CPW resonator circuit (3 dB bandwidth/center frequency), we can obtain $\tan\delta$ of the ferroelectric thin film, as the unloaded Q-factor will be inversely proportional to $\tan\delta$ (i.e., $\tan\delta \approx 1/Q_u$) [23–25].

Figure 16 shows a plot of the ε_{eff} versus $\varepsilon_{r\text{BSTO}}$ for a CPW line with 300 nm BSTO thin film, at 20 GHz. This plot is obtained using Sonnet simulation of a CPW line of 1 cm long, on a 10 mil thick LAO substrate.

Several ferroelectric tunable CPW lines were fabricated on LAO, MgO, and LSAT substrates. One BSTO tunable CPW line (4 mm long) on a 10 mil thick LSAT substrate, with a conductor width of 4 mils and the gap between the conductor and the ground line of 0.8 mils, exhibited a low insertion loss and moderate tunability, as shown in Figure 17. The BSTO film for this sample is approximately 400 nm thick. Note that as the center conductor is biased at a higher voltage level, the dc electric field changes the dielectric constant of the BSTO film. As can be seen from Figure 17, the insertion loss improves as one increases the bias voltage. The $\varepsilon_{r\text{BSTO}}$ as well as $\tan\delta$ both improve at higher bias voltage, resulting in reduced insertion loss. The $\varepsilon_{r\text{BSTO}}$ was extracted from the measurements at zero bias as

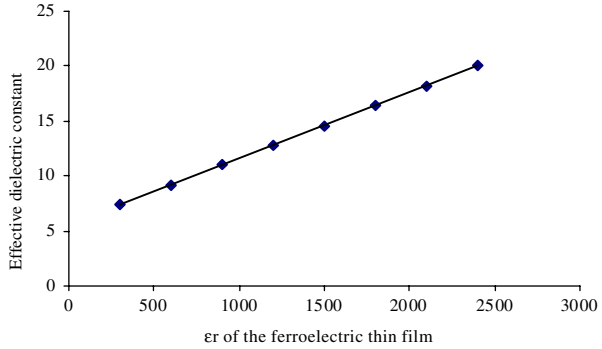


Figure 16. The ε_{eff} vs $\varepsilon_{r\text{BSTO}}$ for a 50 Ω CPW line with a ferroelectric layer thickness of 300 nm, at 20 GHz. The loss tangent of the ferroelectric thin film was assumed to be 0.001 for this simulation.

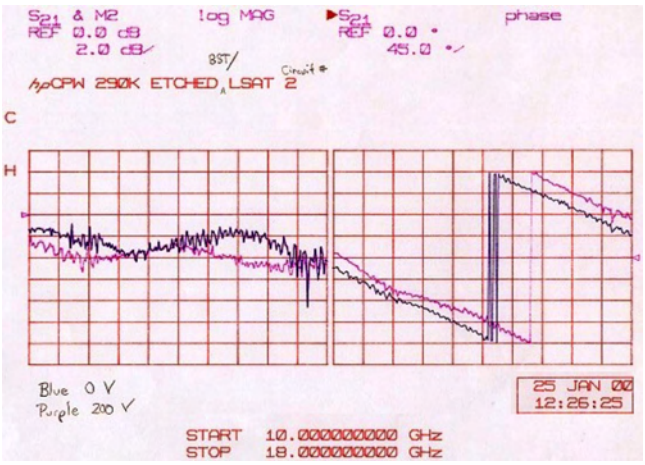


Figure 17. The swept frequency response of magnitude and phase of S_{21} at two different bias voltages applied to the center conductor, in a BSTO (400 nm) tunable CPW line of 4 mm, fabricated on LSAT substrate.

well as at 150 V bias. The measurements agreed with modeling at tuning from $\varepsilon_{r\mathrm{BSTO}}$ of 900 at zero bias to $\varepsilon_{r\mathrm{BSTO}}$ of 300 at 75 kV/cm (corresponding to a dc bias of 150 V). The dielectric tunability for this particular sample is calculated as $[\varepsilon_{r\mathrm{BSTO}}(\text{zero bias}) - \varepsilon_{r\mathrm{BSTO}}(\text{high bias})]/\varepsilon_{r\mathrm{BSTO}}(\text{zero bias})$. The dielectric tunability in this case is approximately 66.67%.

An important figure of merit for tunable microwave devices is the K-factor, defined as the ratio of the phase shift (due to dielectric tunability) to the insertion loss of the sample. The K-factor for the sample in Figure 18 is approximately 30 degrees/dB. The worst-case insertion loss needs to be taken into consideration in the frequency range of the measurements.

An example of the swept frequency measurement for a CPW resonator is shown in Figure 18. One can clearly see the resonance peak from this measurement. The loaded Q of the resonator is calculated as the ratio of the 3 dB bandwidth to the center frequency. In the example, the loaded Q is calculated as 65. The unloaded Q can be obtained from $Q_U = Q_L/[1 - S_{21}(f_0)]$, where $S_{21}(f_0)$ is the ratio of the output voltage to the input voltage as a fraction at the resonant frequency. The unloaded Q is approximately 150 for this example. Tan δ is calculated as the inverse of the unloaded Q. This is based on the assumption that the dielectric loss is primarily due to the ferroelectric thin film layer.

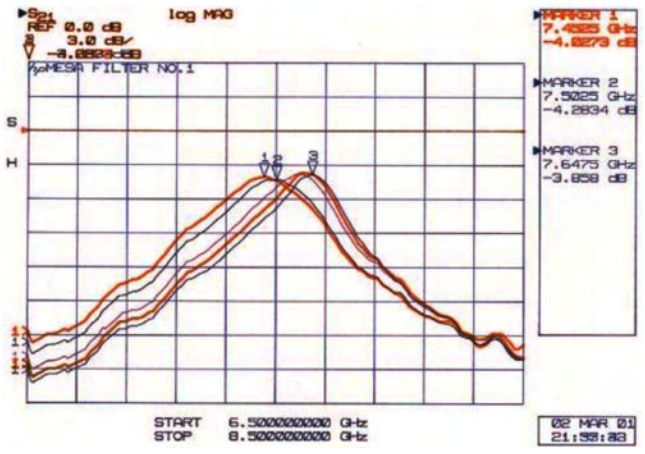


Figure 18. Frequency response of a BSTO tunable resonator showing the improvement in insertion loss as a function of biasing voltage. The left most plot is for the zero-bias voltage, and the right most for +400 V bias on the resonator.

As shown in Figure 18, the center frequency of the resonator is tunable with applied bias, because of the dielectric tunability of BSTO thin film with applied dc bias. In this example, the resonator's passband is very much the same for both 0 and 400 V applied bias. Such a condition is desirable for bandpass filtering applications. The nanophase BST thin films tested to date show that they are tunable by 20%, and the loss tangent of the films is approximately 0.025–0.05 between 1 and 25 GHz. Further studies are underway to optimize the BST thin films to yield lower loss-tangent values.

5. SUMMARY

Utilizing a repeatable and controllable process to synthesize nanocluster and nanostructured materials is essential for understanding the fundamental physics involved in the nanoscience of materials and the evolution of the science into nanotechnology. PLD provides, with established process control, a deposition TECHNIQUE required for the synthesis of high quality multicomponent thin films. As seen in this chapter, there are at least 17 PLD process variables. Most can be monitored and controlled during the synthesis of the nanostructure of a deposited film. During PLD the microstructure of a BSTO films is very sensitive to the ambient chamber pressure. Further tailoring of the microstructure should be enabled by adjusting other deposition parameters individually or combinatorially such as heater temperature or laser repetition rate. However, adjusting (increasing) the heater temperature at a setpoint TOF will change (decrease) the laser beam energy density because the plume will undergo a different amount of (less) collisions before reaching the TOF monitoring control point. An increase in heater temperature decreases the density of the ambient gas in the chamber which translates to less collisions; hence the plume components will move faster at a set energy density. Careful examination of the listed process control variables and their general affect on the critical growth dynamics should enable fabrication of films with the desired characteristics.

Emission spectroscopy permits the examination of plume sensitivity to almost any of the deposition process variables and is a critical element in developing self-directed process control for PLD. XRD reveals information about the lattice structure as well as crystal structure of the film deposited. Another critical tool in the characterization of the film's nanostructure is the SPM. It can be utilized to study physical, chemical, and electrical properties of the film surface.

Although our study on BSTO nanostructured thin films has yet to be completed, engineering the grain/particle size and the domain state of a ferroelectric material is clearly useful to obtain desired microwave properties of the material. The desired result is to produce optimum material characteristics

Table 1. Summary of important studies on bulk and thin-film BSTO.

No.	Topic	Ref.	Description
1	Ferroelectric memories	[107]	High K BST capacitors for DRAM.
2	Temperature coefficient of dielectric constant (TCK)	[108]	Adding small quantities of BSN decreases the TCK dramatically.
3	Giant effective pyroelectric coefficients in graded ferroelectrics	[109]	Pyroelectric coefficients $\sim 5\ \mu C/cm^2\ ^\circ C$ obtained from compositionally graded BST thin films on Si.
4	Effect of stress on dielectric properties of BST thin films	[110]	Stress-free capacitance was found to be 23% higher than the capacitance under residual stress.
5	Dependence of dielectric properties on internal stresses in BST thin films	[111]	BST thin films on MgO substrates, with thickness of 14 to 500 nm, were investigated. Dielectric constant drops from 2350 for highly stressed films to 1700 in relaxed thicker films.
6	Microstructure and dielectric properties of BST films on LAO	[112]	$Ba_{1-x}Sr_xTiO_3$ thin films with $x = 0.1$–0.9 were studied. At room temperature, the dielectric constant and tunability are high for $x \leq 0.4$ and decrease rapidly for as x increases.
7	Infrared optical properties of BST thin films	[113]	$Ba_{0.8}Sr_{0.2}TiO_3$ thin films grown on $Pt/Ti/SiO_2/Si$ substrates were studied in the spectral range of 2.5–12.6 μm. The refractive index decreases and the extinction coefficient increases as the wavelength increases.
8	Pyroelectric properties of sol–gel derived BST thin films	[114]	Pyroelectric coefficient of $Ba_{0.8}Sr_{0.2}TiO_3$ thin films on $Pt/Ti/SiO_2/Si$ substrates. Measured pyroelectric coefficient $> 3.1 \times 10^{-4}$ C/m^2 K from 10 to 26 °C.
9	Dielectric tunability and harmonic generation in BST thin films on LAO	[115]	Dielectric tuning is investigated at nanosecond time-scale in $Ba_{0.3}Sr_{0.7}TiO_3$ thin films at 3 GHz. Full dielectric tuning can be expected at nanosecond time scales.
10	Flexoelectric polarization of BST	[116]	Strain-gradient induced polarization (flexoelectric effect) is studied in $Ba_{0.67}Sr_{0.33}TiO_3$ bulk ceramic in the paraelectric state. The flexoelectric coefficient (μ_{12}) is roughly proportional to dielectric permittivity.
11	BST thin films on Si with Pt electrodes	[117]	$Ba_{0.6}Sr_{0.4}TiO_3$ thin films grown by PLD on $Pt/SiO_2/Si$ substrates. Measured dielectric constant for a 170 nm BST film was 400 with a loss tangent of 0.01–0.03 at 10 kHz.
12	Effect of annealing on the structure and dielectric properties	[118]	The effect of a postdeposition anneal on the structure and dielectric properties of epitaxial $Ba_xSr_{1-x}TiO_3$ thin films ($x = 0.35$–0.65) were measured. The anneal had the effect of an increased electric field dependence on dielectric tuning without increased dielectric loss.
13	Composites of BST and other nonelectrically active ceramics	[25]	Reduced dielectric constant and loss tangent for low-loss tunable dielectric materials.
14	Doped BST for microwave applications at room temperature	[119]	Dielectric properties of laser ablated Mn-doped and undoped $Ba_{0.6}Sr_{0.4}TiO_3$ thin films are compared at low frequencies and at microwave frequencies. 1% Mn doped films had large tunability of 56% at 40 kV/cm and a loss tangent of 0.006 at zero bias and 10 GHz.
15	Nonlinear properties of BST	[120]	Four types of nonlinear properties of ferroelectric materials were considered: paraelectric–ferroelectric interphase boundaries, nonlinear equilibrium excitations–domain walls, nonlinear response of domain structure near the first-order phase transition-wetting of domain walls, and nonlinear effects caused by external high magnetic fields.
16	Strain-induced ferroelectric properties	[121]	Metal–ferroelectric–insulator–semiconductor structures were formed by PLD. Mechanical strain induced a ferroelectricity in the BST films.
17	Enhanced electrical properties	[122]	Enhanced electrical performance of high-dielectric-constant BST thin films grown by *in-situ* UV-assisted PLD in comparison with conventional PLD.
18	BSTO thin films for uncooled infrared focal plane arrays	[123]	BSTO thin films prepared by improved sol–gel technique. BST15 thin film is the advanced candidate material for uncooled infrared focal plane arrays at near ambient temperature.

continued

Table 1. Continued

No.	Topic	Ref.	Description
19	Size effect and fatigue mechanism in ferroelectric thin films	[124]	Fatigue in ferroelectrics may be caused by field screening mechanism where only a fraction of the applied voltage is seen by film bulk.
20	Thickness-dependent dielectric constants of BSTO thin films	[125]	Decrease in measured dielectric constant of BSTO thin films with decreasing dielectric film thickness analyzed.
21	Thin film ferroelectrics for capacitor applications	[126]	PLD used to fabricate thin film capacitor, electrical characterization of BSTO films showed Curie point behavior independent of composition, contrasting sharply with bulk behavior.
22	Structural features and phase transition temperature of BSTO films on various substrates	[127]	BSTO films grown on various substrates by rf sputtering. Structural diagnostics of films correlated with temperature corresponding to maximum capacitance.
23	BST for tunable phase shifters	[128]	The influence of the microstructure of BST: in comparison to dense bulk ceramics, the permittivity of thick films is approximately 10 times less.
24	Advances in low loss, tunable dielectrics	[25]	Composites of BSTO fabricated in bulk ceramic, thick film, and thin film. The dielectric constant and loss tangents of the composites were reduced to enhance impedance matching and lowering overall insertion loss of devices.
25	Microwave power handling	[129]	High power handling capability of ferroelectric thin-film based circuits.
26	Coupled microstrip phase shifters	[130]	BSTO thin-film based CMPS with 60 degrees/dB *K*-factor in *Ku*- and *K*-band phase shifter circuits.
27	Distributed analog phase shifter	[131]	BST varactors periodically loading CPW lines for analog phase shift.
28	Phased array antenna based on bulk phase shifting	[132]	A new concept for phased arrays based on voltage controlled ferroelectric lens implemented using bulk ferroelectric slabs.
29	Electronically steerable reflect array antennas	[133]	Eight element reflect array demonstrated using BST thin-film based CMPS circuits.
30	Ferroelectric/ferrite hybrid circuits	[134]	Combined ferroelectric/ferrite thin films for tunable devices demonstrated.

for use in frequency and phase-agile microwave devices and also gain an understanding of the physics involved in fabricating nanostructured materials. Table 1 is provided as a summary of most important work on BSTO bulk as well as thin films. Further work will continue in our group in optimizing the process for low microwave loss and highly tunable BSTO thin films for room temperature applications.

GLOSSARY

a_t and a_c *a*-axes of the tetragonal and cubic unit cell.

Domain walls Regions between two domains.

Ferroelectric A group of dielectric materials exhibiting spontaneous polarization.

Ferroelectric domain walls Domain walls oriented differently in response to the spontaneous polarization vector.

Ferroelastic domain walls Domain walls which differ in orientation to the spontaneous strain tensor.

Hysteresis loop Consequence of the domain-wall switching in a ferroelectric materials.

Nanoclusters Aggregates of nanoparticles.

Nanostructure Material topography (physical features) consisting of structures of size less than or equal to 10^{-9} meters.

Paraelectric Dielectric material without spontaneous polarization.

Pyroelectric Dielectric displaying polarization without an electric field.

Sol–gel Chemical process for thin film deposition.

Thin film Film with crystal structure, in this case thickness less than 300 nanometers.

Twinning Atoms are located in mirror image positions of atoms on the other side of the boundary.

REFERENCES

1. J. Grindlay, "Introduction to the Phenomenological Theory of Ferroelectricity." Pergamon, London, 1970.
2. D. L. Mills, *Phys. Rev. B* 3, 3887 (1971).
3. D. R. Tilley and B. Zeks, *Solid State Commun.* 49, 823 (1984).
4. R. Castro-Rodriguez, V. E. Quadrelli, F. Calderon, F. Leccabue, and B. E. Watts, *Mater. Lett.* 34, 326 (1998).
5. Y. G. Wang, W. L. Zhong, and P. L. Zhang, *Solid State Commun*. 90, 329 (1994).
6. K. Okuwada, *J. Sol–Gel Sci. Technol.* 16, 77 (1999).
7. A. Gruverman, O. Auciello, and H. Tokumoto, *Integ. Ferr.* 19, 49 (1988).
8. J. Fitz-Gerald, S. Pennycook, H. Gao, and R. K. Singh, *Nanostruct. Mater*. 12, 1167 (1999).
9. H. Shiibashi, H. Matsuda, and M. Kuwabara, *J. Sol–Gel Sci. Technol.* 16, 129 (1999).
10. C. M. Gilmore and J. A. Sprague, *Nanostruct. Mater.* 9, 643 (1997).
11. P. Ayyub, S. Chattopadhyay, K. Sheshadri, and R. Lahiri, *Nanostruct. Mater.* 12, 713 (1999).
12. "Dielectric Materials and Applications" (A. von Hippel, Ed.). Artech House, Boston, 1995.
13. M. Eng, *Nanotechnology* 10, 405 (1999).
14. J. Takasu, *J. Electroceram.* 4, 327 (2000).

15. Y. Tarui, T. Hirai, K. Teramoto, H. Koike, and K. Nagashima, *Appl. Surf. Sci.* 113/114, 656 (1997).
16. "Thin-Film Ferroelectric Materials and Devices" (R. Ramesh, Ed.). Kluwer Academic, Norwell, MA, 1997.
17. G. Subramanyam, F. W. Van Keuls, and F. A. Miranda, *IEEE Microwave Guided Wave Lett.* 8, 78 (1998).
18. F. A. Miranda, G. Subramanyam, F. W. Van Keuls, R. R. Romanofsky, J. D. Warner, and H. Mueller, *IEEE Trans. Microwave Theory Tech.* 48, 1181 (2000).
19. R. R. Romanofsky, J. Bernhard, G. Washington, F. W. Van Keuls, F. A. Miranda, and C. L. Canedy, *IEEE MTT Int. Microwave Symp. Dig.* 2, 1351 (2000).
20. G. Subramanyam, F. Van Keuls, F. A. Miranda, R. Romanofsky, C. Canedy, S. Aggarwal, T. Venkatesan, and R. Ramesh, *IEEE Microwave Guided Wave Lett.* 10, (2000).
21. R. R. Romanofsky, F. W. Van Keuls, and F. A. Miranda, *J. Phys. IV France* 8, 171 (1998).
22. C. H. Mueller and F. A. Miranda, "Ferroelectric and Acoustic Devices" (D. Taylor and M. Francombe, Eds.). Academic Press, San Diego, 2000.
23. E. Carlsson and S. Gevorgian, *IEEE Trans. Microwave Theory Techniques* 47, 1544 (1999).
24. S. Gevorgian S, T. Martinsson, and A. Deleniv, *IEE Microw. Anten.* 144, 145 (1997).
25. F. W. Van Keuls, C. Chevalier, F. A. Miranda, C. M. Carlson, T. V. Rivkin, P. A. Parilla, J. D. Perkins, and D. S. Ginley, *Microwave Opt. Technol. Lett.* 29, 34 (2001); L. Sengupta and S. Sengupta, *Mater. Res. Innovations* l2, 278 (1999).
26. K. Suiz and K. Nagayama, *Appl. Phys. A* 69, S235 (1999).
27. J. Zarzycki, *J. Sol–Gel Sci. Technol.* 8, 17 (1997).
28. X. Y. Chen and Z. G. Liu, *Appl. Phys. A* 69, S5323 (1999).
29. L. Sengupta and S. M. Green, *Appl. Surf. Sci.* 127, 486 (1998).
30. D. Mou, J. Linnros, C. S. Petersson, and K. V. Rao, *J. Appl. Phys.* 84, 5785 (1998).
31. P. T. Han and P. T. Murray, *J. Appl. Phys.* 88, 11184 (2000).
32. M. Ozegowski, K. Meteva, S. Metev, and G. Sepold, *Appl. Surf. Sci.* 138, 68 (1999).
33. I. N. Germanenko, S. Li, S. J. Silvers, and M. S. El-Shall, *Nanostruct. Mater.* 12, 731 (1999).
34. K. Murakami, T. Suzuki, T. Makimura, and M. Tamura, *Appl. Phys. A* 69, S13 (1999).
35. L. Patrone, D. Nelson, V. Safarov, M. Sentis, and W. Marine, *J. Lumin.* 80, 217 (1999).
36. L. Patrone, D. Nelson, V. I. Safarov, S. Giorgio, M. Sentis, and W. Marine, *Appl. Phys. A* 69, S217 (1999).
37. S. Joshi and M. W. Cole, *Appl. Phys. Lett.* 77, 289 (2000).
38. G. Subramanyam, F. W. Van Keuls, F. A. Miranda, C. L. Canedy, S. Aggarwal, T. Venkatesan, and R. Ramesh, *Integrated Ferroelectrics* 24, 273 (1999).
39. M. P. McNeal, S.-J. Jang, and R. E. Newnham, *J. Appl. Phys.* 83, 3288 (1998).
40. L. Zhang, W.-L. Zhong, C.-L. Wang, P.-L. Zhang, and Y.-G. Wang, *J. Phys. D* 32, 546 (1999).
41. B. Jiang and A. Bursill, *Phys. Rev. B* 60, 9978 (1999).
42. C. L. Wang, Y. Xin, X. S. Wang, and W. L. Zhong, *Phys. Rev. B* 62, 11424 (2001).
43. S. Chattopadhyay, *Nanostruct. Mater.* 9, 551 (1997).
44. Y. Wang, W. Zhong, and P. Zhang, *Sci. China* (*Ser. A*) 38, 724 (1995).
45. J. Zarzycki, *J. Sol–Gel Sci. Technol.* 8, 17 (1997).
46. R. M. Almeida, *J. Sol–Gel Sci. Technol.* 13, 51 (1998).
47. H. Shiibashi, H. Matsuda, and Kuwabara, *J. Sol–Gel Sci. Technol.* 16, 129 (1999).
48. H. K. Schmidt, E. Geiter, M. Mennig, H. Krug, C. Becker, and T. P. Winkler, *J. Sol–Gel Sci. Technol.* 13, 397 (1998).
49. H. B. Sharma, H. N. K. Sarma, and A. Mansingh, *J. Appl. Phys.* 85, 341 (1999).
50. T. D. de Morais, F. Chaput, J. P. Boilot, K. Lahlil, B. Darracq, and Y. Levy, *Adv. Mater. Opt. Electron.* 10, 69 (2000).
51. R. N. Viswanath and S. Ramasamy, *Nanostruct. Mater.* 8, 155 (1997).
52. H. Shiibashi, H. Matsuda, and M. Kuwabara, *J. Sol–Gel Sci. Technol.* 16, 129 (1999).
53. "Pulsed Laser Deposition of Thin Films" (D. B. Chrisey and G. K. Hubler, Eds.). Wiley, New York, 1994.
54. P. R. Willmott and J. R. Juber, *Rev. Mod. Phys.* 72, (2000).
55. M. Ozegowski, K. Meteva, S. Metev, and G. Sepold, *Appl. Surf. Sci.* 138, 68 (1999).
56. R. R. Biggers, J. G. Jones, I. Maartense, J. D. Busbee, D. Dempsey, D. Liptak, D. Lubbers, C. Varanasi, and D. Mast, *Eng. Appl. Artificial Intelligence* 11, 627 (1998).
57. F. Stark and S. J. P. Laube, *JMEPEG* 2, 721 (1993).
58. Z. Kantor and T. Szörényi, *Appl. Surf. Sci.* 127–129, 703 (1998).
59. R. K. Singh, *J. Non-Cryst. Solids* 178, 199 (1994).
60. J.-K. Heinsohn, D. Reimer, A. Richter, K.-O. Subke, and M. Schilling, *Physica C* 299, 99 (1998).
61. T. E. Itina, W. Marine, and M. Aurric, *Appl. Surf. Sci.* 154–155, 60 (2000).
62. P. Chu, P. P. Ong, H. f. Chen, and H. H. Teo, *Appl. Surf. Sci.* 137, 91 (1999).
63. T. Svobodny and R. R. Biggers, *Proc. SPIE* 4760 (2002).
64. Z. Trananovic, I. Takeuchi, P. A. Warburton, C. J. Lobb, and T. Venkatesan, *Physica C* 265, 79 (1996).
65. D.-Y. Kim, S.-G. Lee, Y. K. Park, and S. J. Park, *Mater. Lett.* 40, 146 (1999).
66. M. Mukaida, S. Sato, Y. Takano, M. Kusunoki, and S. Ohshima, *Physica C* 378–381, 1232 (2002).
67. R. Biggers, G. Norton, and I. Maartense, private communication.
68. R. Biggers, G. Kozlowski, J. Jones, D. Depsey, R. Kleismit, I. Martense, J. Busbee, T. Peterson, and R. Perrin, *Integrated Ferroelectrics* 28, 201 (2000).
69. D. B. Geohegan and A. A. Puretzky, *Appl. Phys. Lett*. 67, 197 (1995).
70. C. E. Otis and R. W. Dreyfus, *Phys. Rev. Lett.* 67, 2101 (1991).
71. A. P. McKierman and J.-P. Mosnier, *Appl. Surf. Sci.* 197–198, 325 (2002).
72. R. Biggers, G. Kozlowski, J. Jones, D. Dempsey, R. Kleismit, I. Maartense, J. Busbee, and T. Peterson, *Integrated Ferroelectrics* 28, 201 (2000).
73. W. Marine, L. Patrone, B. Luk'yanchuk, and M. Sentis, *Appl. Surf. Sci.* 154–155, 345 (2000).
74. "Clusters of Atoms and Molecules" (H. Haberland, Ed.). Springer-Verlag, Berlin, 1994.
75. T. Yoshida, Y. S. Takeyama, Y. Yamada, and K. Mutoh, *Appl. Phys. Lett.* 68, 1772 (1996).
76. J. Gonzalo, R. Serna, J. M. Requejo, C. N. Afonso, and A. Naudon, *Appl. Surf. Sci.* 154–155, 449 (2000).
77. Z. Pászti, G. Petô, Z. E. Horváth, and A. Karacs, *Appl. Surf. Sci.* 168, 114 (2000).
78. P. Periaswamy, M. J. Lefevre, T. R. Taylor, and R. York, *Appl. Phys. Lett.* 75, 3186 (1999).
79. M.-C. Wang, C. C. Tsai, and N. C. Wu, *J. Appl. Phys.* 92, 2100 (2002).
80. Y. Gao, P. Alluri, and S. He, *J. Appl. Phys.* 87, 124 (2000).
81. A. Tombak, J. P. Maria, F. Ayguavives, Z. Jin, G. T. Stauf, A. I. Kingon, and A. Mortazawi, *IEEE Microwave Wireless Components Lett.* 12, 3 (2002).
82. Y.-A. Jeon, E.-S. Choi, T.-S. Seo, and S.-G. Yoon, *Appl. Phys. Lett.* 79, 1012 (2001).
83. T. Zhao, F. Chen, H. Lu, G. Yang, and Z. Chen, *J. Appl. Phys.* 87, (2000).
84. M. P. McNeal, S.-J. Jang, and R. E. Newnham, *J. Appl. Phys.* 83, 3288 (1998).

85. D. Y. Kim, S. G. Lee, Y. K. Park, and S. J. Park, *Mater. Lett.* 40, 46 (1999).
86. S. Vijayalakshmi, M. A. George, J. Sturmann, and H. Grebel, *Appl. Surf. Sci.* 127–129, 378 (1998).
87. H. Jaschke, H.-J. Butt, S. Srin, H. E. Gaub, O. Hasemann, F. Krimphove, and E. K. Wolff, *Biosensors Bioelectron.* 11, 601 (1996).
88. Scanning Probe/Atomic Force Microscopy: Technology Overview and Update, Application Note from Digital Instruments website, http://www.veeco.com/html/product_bymarket.asp.
89. C. J. Maal, *Res. Develop. Mag.* 39, 22 (1997).
90. K. Jandt, *Mater. Sci. Eng.* R21, 221 (1998).
91. The STM—An introduction, http://www.cmmp.ucl.ac.uk/~atg/work/stm.html.
92. SEM and AFM: Complementary Techniques for High Resolution Surface Investigations, Application Note from Digital Instruments website, http://www.veeco.com/html/product_bymarket.asp.
93. Digital Instruments, Veeco Metrology group, Multi-Mode SPM Manual, version 4.31ce.
94. D. H. Lowndes, C. M. Rouleau, T. Thunday, G. Duscher, E. A. Kenik, and S. J. Pennycook, *Appl. Surf. Sci.* 127–129, 355 (1998).
95. T. Yoshida, Y. Yamada, and T. Orii, *J. Appl. Phys.* 83, 5427 (1998).
96. Makimura, Y. Kunii, and K. Murakami, *J. Appl. Phys.* 35, 4780 (1996).
97. W. Chang and L. Sengupta, *J. Appl. Phys.* 92, 3941 (2002).
98. L. Sengupta, E. Ngo, S. Stowell, M. O'Day, and R. Lancto, U.S. Patent 5, 427, 988, 1995.
99. K. S. Grabowski, J. S. Horwitz, and D. B. Chrisey, *Ferroelectrics* 116, 19 (1991).
100. F. Zimmermann, M. Voigts, C. Weil, R. Jakoby, P. Wang, W. Menesklou, and E. Ivers-Tiffee, *J. Eur. Ceram. Soc.* 21, 2019 (2001).
101. M. A. Delfrate, M. Leoni, L. Nanni, E. Melioli, B. E. Watts, and F. Leccabue, *J. Mater. Sci. Mater. Electron.* 5, 153 (1994).
102. L. C. Sengupta, S. Stowell, E. Ngo, M. E. O'Day, and R. Lancto, *Integragted Ferroelectrics* 8, 77 (1995).
103. A. J. Moulson and A. J. *British Ceram. Proc.* 36, 57 (1985).
104. G. Arlt, D. Nennings, and D. J. De With, *J. Appl. Phys.* 58, 71 (1985).
105. M. Jain, S. B. Majumder, R. S. Katiyar, D. C. Agrawal, and A. Bhalla, *Appl. Phys. Lett.* 81, (2002).
106. G. Subramanyam, N. Mohsina, A. Zaman, F. W. Van Keuls, F. A. Miranda, R. R. Romanofsky, J. D. Warner, and C. L. Chen, *Integrated Ferroelectrics* 42, 151 (2002).
107. T. Eimori et al., *IEDM Dig.* 631 (1993).
108. R. J. Cava, W. F. Peck, J. J. Krajewski, and D. A. Fleming, *Appl. Phys. Lett.* 67, 3813 (1995).
109. F. Jin et al., *Appl. Phys. Lett.* 73, 1838 (1999).
110. T. M. Shaw, Z. Suo, E. Liniger, M. Huang, and R. B. Laibowitz, *Appl. Phys. Lett.* 75, 2129 (1999).
111. H. Li, A. L. Roytburd, S. P. Alpay, T. D. Tran, L. Salamanca-Riba, and R. Ramesh, *Appl. Phys. Lett.* 78, 2354 (2001).
112. Y. Gim, T. Hudson, Y. Fan, C. Kwon, A. T. Findikoglu, B. J. Givvons, B. H. Park, and Q. X. Jia, *Appl. Phys. Lett.* 77, 1200 (2000).
113. Z. Huang, Z. Zhang, C. Jiang, J. Yu, J. Sun, and J. Chu, *Appl. Phys. Lett.* 77, 3651 (2000).
114. J.-G. Cheng, J. Tang, J.-H. Chu, and A.-J. Zhang, *Appl. Phys. Lett.* 77, 1035 (2000).
115. J. C. Booth, R. H. Ono, I. Takeuchi, and K.-S. Chang, *Appl. Phys. Lett.* 81, 718 (2002).
116. W. Ma and L. E. Cross, *Appl. Phys. Lett.* 81, 3440 (2002).
117. L. Kinder, X. F. Zhang, I. L. Grigorov, C. Kwon, Q. x. Jia, L. Luo, and J. Zhao, *J. Vac. Sci. Technol. A* 17, 2148 (1999).
118. L. K. Knauss, J. M. Pond, J. S. Horwitz, D. B. Chrisey, C. H. Mueller, and R. Treece, *Appl. Phys. Lett.* 69, 25 (1996).
119. H. Wu and F. Barnes, *Integrated Ferroelectrics.* 22, 811 (1998).
120. S. Dorfman, D. Fuks, A. Fordon, E. Kotomin, and P. Wyder, *Physica B* 304, 339 (2001).
121. S. Jun, Y. K. Kim, and J. Lee, *Surf. Coatings Technol.* 131, 553 (2000).
122. A. Srivastave, V. Craciun, J. Howard, M. Joshua, and R. K. Singh, *Appl. Phys. Lett.* 75, 3002 (1999).
123. S. Liu, M. Liu, Y. Zeng, C. Li, S. Chen, Y. Huang, and D. Xia, *Mater. Sci. Eng. C* 22, 73 (2002).
124. H. Z. Jin and J. Zhu, *J. Appl. Phys.* 92, 4594 (2002).
125. C. S. Hwang, *J. Appl. Phys.* 92, 432 (2002).
126. D. O'Neill, G. Catalan, F. Porras, R. M. Bowman, and J. M. Gregg, *J. Mater. Sci. Mater. Electron.* 9, 199 (1998).
127. S. F. Karmanenko, A. I. Dedyk, A. A. Melkov, R. N. Il'in, V. I. Sakharov, I. T. Serenkov, and H. Jowoong, *J. Phys. Condensed Matter* 14, 6823 (2002).
128. F. Zimmermann, M. Voigts, C. Weil, R. Jakoby, P. Wang, W. Memesklou, and E. Ivers-Tiffee, *J. Eur. Ceram. Soc.* 21, 2019 (2001).
129. D. Galt, T. Rivkin, and M. Cromar, *Mater. Res. Soc. Symp. Proc.* 493, 341 (1998).
130. F. W. Van Keuls, R. Romanofsky, N. D. Varaljay, F. A. Miranda, C. L. Canedy, S. Aggarwal, T. Venkatesan, and R. Ramesh, *Microwave Opt. Technol. Lett.* 20, 53 (1999).
131. Y. Liu, A. Nagra, E. G. Erker, P. Periaswamy, T. R. Taylor, J. Speck, and R. A. York, *IEEE Microwave Guided Wave Lett.* 10, 448 (2000).
132. JBL Rao, D. P. Patel, and L. Sengupta, *Integrated Ferroelectrics* 22, 827 (1998).
133. R. R. Romanofsky, J. T. Bernhard, F. W. Van Keuls, F. A. Miranda, G. Washington, and C. Canedy, *IEEE Trans. Microwave Theory Technol.* 48, 2504 (2000).
134. H. Jiang, W. Hu, S. Liang, V. Fouflyguine, J. Zhao, Q. X. Jia, J. R. Groves, P. Arendt, F. A. Miranda, A. Drehman, S. Wang, and P. Yip, *Integrated Ferroelectrics* 29, 63 (2000).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanostructured Biomaterials

R. Murugan, S. Ramakrishna

National University of Singapore, Singapore

CONTENTS

1. INTRODUCTION

Recent years have clearly shown a considerable upsurge in the number of publications concerning biomaterials engineered at the nanolevel. The field of biomaterials, a multidisciplinary subject, has developed rapidly due to its unique potential in human health care systems. According to the definition of Williams [225], a biomaterial is a substance or a combination of substances, other than drugs, derived either naturally or synthetically used to repair or replace the function of living tissues of the human. An ideal biomaterial should be highly biocompatible, bioactive, easy to fabricate, reproducible, sterilizable, nontoxic, noncorrosive, nonimmunogenic, and not cause any foreign body reactions upon implantation. The use of biomaterials in humans dates far back into ancient civilizations as Egyptians used artificial eyes, ears, teeth, and noses on mummies [224]. They used elephant's tusks, walrus teeth, and some kinds of wood to replace bone or missing teeth. As per historical evidence, the Indians and Chinese used waxes, glues, and tissues in reconstructing defective parts of the human body. Over the centuries, advancements in synthetic materials and surgical techniques have permitted the use of biomaterials in many ways to restore, maintain, or improve the functions of human tissues or organs [172].

More recently, biomaterials engineered at the nanometer scale have gained recognition in various biomedical applications owing to their potential advantages over the conventional biomaterials in terms of superior surface characteristic properties, high degree of monodispersity, flexibility to design improved biomaterials, reproducibility, and close surface contact with body environments due to a larger surface-area-to-volume ratio. A unique value of nanomaterials is their extremely high surface area having many more sites for achieving property enhancements, which makes them ideal for a wide variety of applications. Moreover, and perhaps more importantly, it provides the opportunity for the user to tailor make the desired properties of such materials according to specific needs through chemical and/or physical modifications. Nanotechnology is mainly concerned with the design and manufacture of materials to nanometric tolerances, a nanometer being equal to 10^{-9} m. Dimensions of some of the materials are given in Figure 1, with the highlights of nanometer-scale biological systems. The field of nanoscale biomaterials is an exciting multidisciplinary subject covering many areas of modern science, engineering, and medicine. The interactions between the various disciplines are where much of the novel activity resides. The scientific literature clearly shows a significant increase in the number of researchers from various disciplines entering into the nanobiomaterials field each year. This results in an increasing number of novel ideas and exciting new opportunities, which in turn impact global nanostructural science and technology. This implies the potentiality of nanostructured biomaterials to address the medical challenges facing day-to-day life.

This review is mainly intended to provide an overview of nanostructured biomaterials, in particular, ceramics and polymers, used in bone grafting, tissue engineering, and drug delivery systems. The chapter also presents the critical issues and scientific challenges that need further research and development of the nanostructured biomaterials for their enhanced acceptance in human health care. For the benefit of readers, fabrication methodology, properties, various types, and the potential of currently available biomaterials are briefly described in each section, in addition to their performance with living tissues or organs. This review, with hope, could be useful for gaining insight and knowledge on nanoscale biomaterials, and opens new avenues for further extension of research in novel and challenging areas of human health-care systems.

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (595–613)

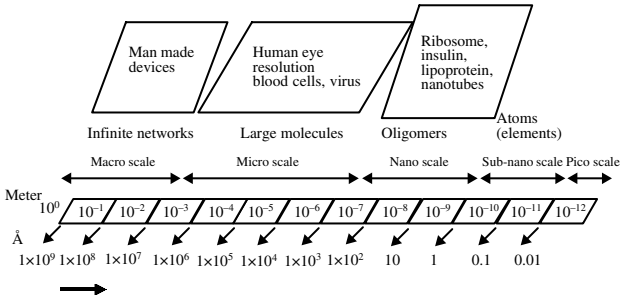


Figure 1. Dimensions of various materials.

2. NANOMATERIALS IN BONE GRAFTING

The need to transplant bone to a deficient skeletal location has been recognized for centuries. Skeletal deficiencies from trauma, tumors, bone diseases, or abnormal development frequently require surgical procedure to restore normal bone functions. If a fracture is minor, usually bone has the capability to self-regenerate its functions in a few weeks' time. In the case of a severe fracture, a bone-grafting substitute may be needed to restore their routine functions. The terminology of bone transplantation is more complicated by the various histological types of bone grafts. Bone-grafting techniques can be grouped into four different types with respect to grafting terminology and implant materials used, namely, autograft, allograft, xenograft, and alloplast. A typical diagram of bone grafting is shown in Figure 2.

Autograft or autogenous graft is the ultimate method being considered as the gold standard as tissue transplanted from one site to another within the same individual. The concept is basically that our own bone is taken from a donor site and placed somewhere else in the same body into the recipient site. The best success rate in bone grafting has been achieved with autografts because these are essentially living tissues with their cells intact. Clinically, it has been proven that there is no chance of immune reaction using this method, and the microscopic architecture is perfectly matched with the host tissue. The only disadvantage of the autograft is that it has to be harvested from a secondary site in the body, which means more morbidity and a more complicated surgery, that is, secondary surgery is needed. For most grafting purposes confined to oral implantology, we can use another part of the jaw (i.e., chin or back portions of the jaw) as an acceptable donor site. This way, we stay surgically inside the mouth, and avoid any extra oral wounds and scarring. Sometimes, however, when there is no availability of bone volume intraorally, it is necessary to get bone from other parts of the body, for example, the hip bone or tibia (shin) bone, since these are the most accessible areas to obtain larger quantities of bone. The second method is allograft or allogenic graft, which is defined as a tissue transplant between individuals of the same species, but of nonidentical genetic composition. This method eliminates the need for a secondary surgery, but after sterilization, the donated bone loses much of its strength. In addition, when bone grafts are taken from cadavers, there is a risk of rejection or of transmitted diseases such as hepatitis B or AIDS caused by infection with the human immunodeficiency virus. Allograft procedures have been on the decline since late 1993 when the U.S. Food and Drug Administration (U.S. FDA) issued regulations governing bone and tissue banks because of the risk of transmitting infectious diseases. The third method is xenograft or xenogenic graft, which pertains to tissue transplant between two different species (e.g., bone from animal to human). Tissue banks usually stock these kinds of materials because it is possible to extract a larger amount of bone with a specific microstructure as compared to bone from human origin. However, there is a possibility of infectious disease transformation from animal bone to implant site if we use this method. Another option, which is becoming more viable, is the use of synthetic materials (biomaterials) to repair or to replace lost bone, known as alloplast or alloplastic graft. These biomaterials have the advantage of eliminating the need for surgery to claim bone for the graft procedure, and eliminating the chance for rejection or transmission of infectious disease. An additional benefit of using this kind of material lies in their availability, reduction of medical cost, and faster recovery time.

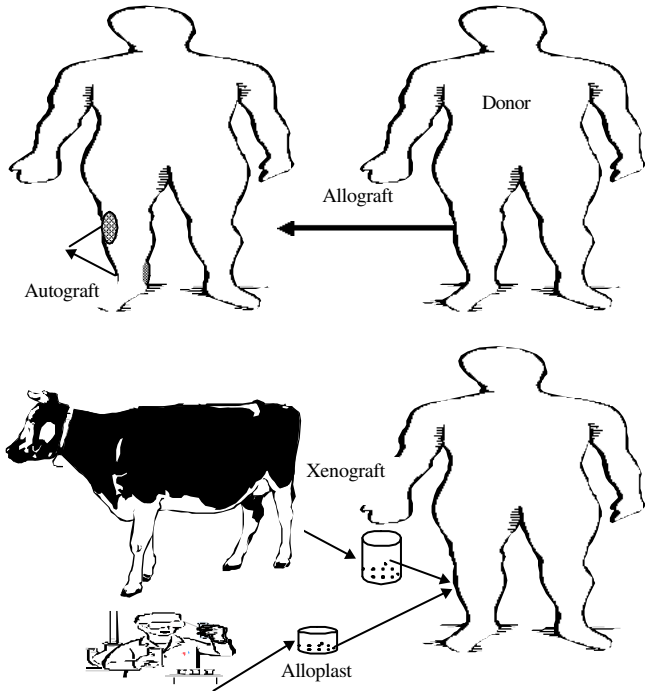


Figure 2. Schematic representation of bone grafting.

Several biomaterials have been developed for bone substitution in the past three decades, and have been used clinically [111, 209, 44, 9]. The major classifications of biomaterials are listed in Table 1 with suitable examples. The materials commonly used as bone substitutes are typical metals, ceramics, polymers, and their composites. Mostly, metals and ceramics are used in hard-tissue applications and polymers in soft-tissue applications due to their structural and mechanical properties. Among them, ceramics and polymer-based biomaterials play a key role in producing artificial bone-graft substitutes. In this topic, we address the potential of those materials in bone grafting and bone tissue engineering applications. The most commonly used

Table 1. Broad classification of biomaterials pertaining to their properties.

Biomaterials	Properties	Examples
Metals	Corrosive	Ferrous
	Noncorrosive	Stainless steel, titanium
Ceramics	Bioinert	Alumina, zirconia
	Bioactive	Hydroxyapatite, bioglass
	Bioresorbable	Tricalcium phosphate
Polymers	Biodegradable	Collagen, gelatin, PLA, PGA, PLGA, PCL
	Nonbiodegradable	PMMA, PHEMA, PEEK, PAA

ceramics and polymers in bone-graft systems include alumina, zirconia, tricalcium phosphates (TCPs), hydroxyapatite (HA), calcium carbonate, dental plaster, demineralized bone matrix, collagen, polylactic acid (PLA), polyglycolic acid (PGA), poly(lactide-co-glycolide) (PLGA), polyethylene glycol (PEG), and polycarbolactone (PCL) [36, 41, 18, 221, 186, 208], which are explained in detailed fashion in this review. The design of medical implants with surface properties similar to physiological bone, characterized by surface grain sizes on the nanometer scale, would undoubtedly help in the formation of new bone at the biomaterial–tissue interface [116]. It is possible to synthesize such materials with nanostructural characteristic features for medical applications with the aid of nanotechnology in order to mimic the natural bone [192].

2.1. Nanobioceramics

Recently, then has been an increasing demand for biomaterials at the nanolevel to assist in the repair of osseous defects caused by skeletal deficiencies in which a substantial amount of bone has been lost. Nanoscale biomaterials, in particular, bioceramics, play a pivotal role as a bone-graft substitute due to their large surface area and their particle size which resembles that of natural components of the bone. In general, ceramics used for the repair and reconstruction of diseased or damaged parts of the human organs or tissues are termed bioceramics. Bioceramics have evolved to become an integral and vital figment of today's modern health-care systems. They are the natural choice of materials with a broad range of composition. The bioceramics used in medicine are primarily classified into three groups: bioinert (e.g., alumina and zirconia), bioactive (e.g., hydroxyapatite), and bioresorbable (e.g., tricalcium phosphates). These classified bioceramics and their performance with respect to tissue reaction are briefly shown in Table 2 [49]. Often, bioceramic materials are used in hard-tissue applications (less load bearing) due to their favorable mechanical properties (Table 3). Mostly, grain sizes of those bioceramics are on the order of few microns with respect to specific site applications. More recently, the term known as nanobioceramics frequently appears in articles in almost all of the fields, particularly biomedical. This may be defined as the ceramics intended for medical applications whose particles in the range of the nanometer scale are known as nanobioceramics. It has unique advantages by having a high surface area and ultrafine structure, which are more

Table 2. Types of bioceramics with respect to their tissue reaction.

Type of bioceramics	Description of biological performance	Examples
Bioinert	Coexistence with implant without noticeable changes and separation from the implant by a layer of fibrous tissue	Single crystals and poly crystalline alumina, zirconia, titanium
Bioactive	Attachment with tissue by direct biochemical bonding, which allows bone-free growth	Hydroxyapatite, bioactive glasses
Bioresorbable	Material is gradually replaced by new bone with respect to their dissolution without any sign of toxicity	Porous hydroxyapatite, tricalcium phosphates, calcium sulfate

beneficial for cell–biomaterial interaction upon implantation. With advances in ceramics and nanotechnology, the application of nanobioceramics as bone substitute has recently received considerable attention because they are biocompatible, bioactive (i.e., direct chemical bonding with bone to biomaterial), provoke little, if any, inflammatory response, and have a large surface area [111, 116, 56, 131].

In general, nanoceramic material synthesis can be grouped into two categories, namely, gas phase and sol–gel process. The gas-phase condensation process was invented by Siegel [193, 194]. This process involves the production of physical vapor from elemental or reacted material, followed by sudden condensation and reaction of the vapor into small nanometer-scale particles. The gas-phase condensation process can be explored in the synthesis of inorganic and metal phases [20]. The sol–gel technique is a wet-chemical approach that can be used to generate nanoparticles by gelation, precipitation, and hydrothermal treatments [123]. An example of the sol–gel process is described by Wiltzius [226]. Here, a concentrated dispersion of colloid is chemically converted into a gel, and then dried, followed by sintering, which produced a ceramic or glass product. By this technique, it is possible to produce nanosize particles, fibers, films, plates, and tubes at low temperature. It is reported that nanoparticles with diameters ranging from 1 to 10 nm with consistent crystal structure, surface derivation, and a high degree of monodispersity have been processed by both the gas-phase and sol–gel techniques [155].

2.1.1. Biocompatible Alumina

Alumina (Al_2O_3) is the first of its kind of bioceramic material widely used clinically [93]. It is frequently being used in total hip prostheses and dental implants due to its multiple characteristic features, including excellent corrosion resistance, biocompatibility, being chemically inert, having low friction, high wear resistance, and high strength [167]. Alumina with an average grain size of less than 4 μm and a purity of greater than 99.7% exhibits good flexural strength and excellent compressive strength. Moreover, this kind of nanometer-scale medical-grade alumina certainly has a higher surface area for better interaction with host tissue.

Table 3. Properties of bioceramics favorable for biomedical applications [33].

Bioceramics	Young's modulus (GPa)	Tensile strength (GPa)	Compressive strength (GPa)	Fracture toughness (MPa · $m^{1/2}$)	Hardness (HV)	Applications
Alumina	366–380	0.31	3.3.8	4	2200	Orthopedic and dental implants, maxillofacial reconstruction
Zirconia	220	0.82	2	7–10	1200	Orthopedic, biocomposites
Hydroxyapatite	80–110	0.05	0.4–0.9	0.7–1.2	600	Coatings on bioimplants, bone and dental filler, periodontal pockets, alveolar ridge augmentation, maxillofacial reconstruction, tissue engineering

Nanoscale alumina is also used in corneal replacements, alveolar ridge and maxillofacial reconstructions, segmental bone replacements, ossicular bone substitutes, knee prostheses, and post-type dental implants [91, 206, 217]. Boutin et al. [19] reported that the alumina as a component of total hip prostheses outperformed other orthopedic implants such as polyethylene. They stated that the long-term success in orthopedic applications depends on its micro/nanostructure, implant geometry, and the health status of the host bone. A successful investigation of osteogenic cell–substrate interaction using nanoscale alumina was reported [217]. Webster et al. prepared nanoalumina in combination with titanium oxide (TiO_3) by compacting with different nanophase grain sizes. They followed the sterilization method for the sample preparation as similarly reported by Puleo et al. [179]. They studied the efficacy of the materials with osteoblast cells, a well-known cell pertaining to the formation of bone, and found that there was an increasing trend in osteoblast adhesion on the nanostructure of the alumina phase. This kind of investigation provided evidence that the enhanced adhesion is independent of the specific chemistry of material surface, and dependent only on the optimal surface topography of nanophase bioceramics.

The extent of osteointegration between bone and a newly implanted material is influenced by many factors, including surface topography. The surface properties of the biomaterials play a vital role in controlling the type and magnitude of cellular and molecular events at the tissue–implant interface. Hence, control and/or manipulation of the surface properties of nanoparticles are of great importance when intending the nanoceramics as medical implants. In general, there are two ways to modify the surface of nanoparticles. The first method is carried out through surface agents by a variety of functionalities depending upon the site of application [213]. The second one is based on grafting polymeric molecules through covalent bonding to the functional groups existing on the nanoparticles [120]. The advantage of the second method over the first lies in the fact that the polymer-grafted particles can be provided with the desired functional properties through a proper section of monomers for the grafting and standardizing optimal grafting condition. In this way, the structure–property relationship of the nanomaterials can be tailor made to a specific need [185]. Recently, grafting of vinyl monomers such as polystyrene and polyacrylamide onto nanometer-scale alumina was successfully carried out [184]. To enhance the interfacial interaction in alumina nanoparticles, the alumina surface was first treated with silane, followed by radical graft polymerization in aqueous and nonaqueous media using a peroxide initiator for an aqueous medium and an azobis–isobutyronitrile (AIBN) initiator for a nonaqueous medium. They tailored the structural properties of the grafts by changing reaction conditions, and found that the grafting polymer radicals and/or the grafted polymer chains had a blocking effect on the diffusion of radicals or monomers toward the surface of alumina nanoparticles. This can be because the interactions between solvent and grafted polymers were weaker as compared to the grafted polymers and the nanoparticles. Another class of materials known as nanofibers also has attracted great attention from materials researchers because of their superior properties and their potential for a wide range of applications. Recently, alumina nanofibers were synthesized at a moderate condition in an aqueous medium through a surfactant-directed crystal growth process using polyethylene oxide (PEO), which is different from the conventional crystallization and the template synthesis of nanofibers thus far studied [236, 237]. This kind of nanofiber is very useful in developing nanocomposites with desired morphology, tailorable size, porosity, and surface topography.

2.1.2. Bioinert Zirconia

Zirconia (ZrO_2), a well-known ceramic oxide, was identified as such in 1789 by the German chemist Klaproth in the reaction product obtained after heating some kind of gems, and was used for a long time blended with rare-earth oxides as a pigment for ceramics. The research and development on zirconia as a biomaterial were started only in the late 1960s, and was studied extensively by Helmer and Driskell [92]. The zirconia, either of pure form or of partially stabilized zirconia, was used for the fabrication of ball heads in total hip replacement (THR), and also recommended for various dental and orthopedic applications [31, 160]. The

basic requirements to use zirconia as medical implants are described by the International Standard Organization (ISO) [106]. More than 3,00,000 zirconia, namely, tetragonal polymorph, ball heads have been implanted [30], and only two failure cases have been reported so far [99].

In this connection, biocompatibility is one of the important parameters for the materials when considering implantation. Zirconia is a biocompatible material, but does not have any direct bonding to bone [167, 83]. There are many reports available for the biocompatibility of zirconia using different cell lines [88]. It is reported that zirconia has no cytotoxic effects when fibroblasts are cultured, and the adhesion and spreading of cells were evaluated [45, 128, 24]. On the other hand, higher inhibition of cellular growth using zirconia and tricalcium phosphate in comparison to alumina or hydroxyapatite was observed in the direct contact test, cell viability, and MTT assay [192]. Ito et al. [108] studied the cytotoxic effects on L929 fibroblasts cocultured in the presence of pseudoextracellular fluid used as a lubricant, and reported that the cytotoxicity of wear products obtained by the friction of ultramolecular-weight polyethylene (UHMWPE) against yittria partially stabilized zirconia (YPSZ) was higher than the cytotoxicity of UHMWPE/Ti-6Al-4V wear products. They found that the cytocompatibility of wear products was improved using a specific pore size filter (0.22 μm). Since toxic effects could be related neither to ion release nor to polyethylene particles less than 0.22 μm in diameter, the presence of floats on the surface of the culture medium may be considered as one of the causes of the cytotoxicity. The effect of wear products of the alumina and zirconia ceramics were evaluated using flow cytofluorimetry analysis [27]. Alumina and zirconia ceramics were tested in comparison with high-density polyethylene on macrophages J774 and RAW267.5, and it was found that polyethylene particles were more cytotoxic than ceramic particles, even if a quantity of polyethylene less than ceramic particles was phagocytosed by macrophages. These results further supported the biocompatibility of zirconia. There are many direct evidences also available for the *in vivo* biocompatibility test of zirconia implanted in bone and soft tissues. The first *in vivo* experiment onto dense zirconia in combination with Y_2O_3 was performed in a monkey's femur, and the absence of adverse tissue reactions with surrounding environments was reported [92]. Later, Hulbert et al. [97] reported the effects of disks and tubes of different composition of nanopore zirconia particles into rabbit muscles. Histological analysis of all of the implants tested nine months after implantation did not show any sign of toxic, immune, or carcinogenic effects. Another study on the biocompatibility of zirconia was performed by Bortz and Onesto [17], who introduced zirconia tubular elements into rabbits and dogs in order to develop tracheal prostheses. They observed the problems of heavy fibrous growth within the lumen of the tube, which sometimes closed it off. There was a sign of bacterial infections observed in rabbits, but never in dogs; nevertheless, prostheses were firmly attached to the trachea, even though infection was present. The *in vivo* implantation of magnesia partially stabilized zirconia (Mg–PSZ) in the paraspinal muscle of rabbits was also investigated and the biocompatibility at different time intervals of postimplantation was evaluated [71]. There was no evidence of any adverse tissue reactions using zirconia as the implant material. In an overall view, the results obtained so far suggest that the zirconia is a highly biocompatible ceramic material due to the evidence of connective tissue formation at the bone–ceramic interface.

2.1.3. Bioactive Nanohydroxyapatite

The development of calcium-phosphate-based ceramics in medical applications started almost three decades ago with the first successful publication of Levitt et al. [127]. Since then, calcium phosphates have gained much interest in biomaterials research due to their physicochemical properties as evidenced by the large number publications until today. Hydroxyapatite is a major type of calcium phosphate bioceramics, with a chemical formula of $Ca_{10}(PO_4)_6(OH)_2$, which is primarily being used in hard-tissue repair and regeneration. It is highly biocompatible, bioactive, osteoconductive, nontoxic, nonimmunogenic, has a direct chemical bonding to bone, and a structure crystallographically similar to bone mineral with an adequate amount of carbonate substitution [5]. The physicochemical properties of HA have wide biological and industrial applications [50, 87, 227, 125, 94]. Basically, it is a family of apatite phase, and the name apatite was first coined by Werner in 1786 [220]. The name was derived from the Greek word for "deceive" because it had been confused with several other minerals of similar appearance, including beryl, amethyst, olivine, and fluorite. It had been suspected earlier by some workers to be a distinct species, but its color and habit variations are distinct from similar appearing minerals. Chemical analysis in 1785 by Proust and Klaprota established apatite as a calcium phosphate, and confirmed its similarity to calcined bone, although the role of the OH, F, and Cl was not recognized at that time [196]. The name "apatite" belongs to a family of components having a hexagonal system of space group $P6_3/m$ (sixfold symmetry axis with a threefold helix and a microplane) with lattice cell dimensions $a = b = 9.42$ Å and $c = 6.88$ Å [42, 118]. The first synthesis of apatite was carried out by Daubree in 1851 [39], who obtained HA by passing phosphorus trichloride vapor over red hot lime. Since then, a number of preparation methods of HA have been reported and reviewed extensively [109]. There is a long history of preparation of HA in the micrometer scale range in various forms like powder, dense, porous, or composite for biomedical applications [93]. In the past few years, significant research efforts have been devoted to the nanostructure processing of HA and its composites as a means to achieve an ultrafine structure with adequate physicochemical and biological properties resembling those of natural bone minerals [214, 215, 37, 48]. Some of the physiochemical properties of HA are listed in Table 4. In general, the routes of HA preparation are mainly grouped as solid state, wet chemical, hydrothermal, mechanochemical, and more recently, microwave processing [126, 115, 110, 150, 151, 94, 152, 90, 161]. Solid-state reaction has generally been used for the processing of ceramic powders at high temperature. The powders prepared by this method usually have irregular shapes with a larger grain size, and often exhibit heterogeneity in the phase composition due to the informporate reaction resulting from small diffusion coefficients of

Table 4. Physical, mechanical, and chemical properties of hydroxyapatite [13, 42, 93, 126].

Properties	Measurement data
Crystal system	Hexagonal
Space group	$P6_3/m$
Cell dimensions (Å)	$a = b = 9.42$, $c = 6.88$
Young's modulus (GPa)	80–110
Elastic modulus (GPa)	114
Compressive strength (MPa)	400–900
Bending strength (MPa)	115–200
Density (g/cm^3)	3.16
Relative density (%)	95–99.5
Fracture toughness ($MPa \cdot m^{1/2}$)	0.7–1.2
Vickers harness	600
Decomposition temp. (°C)	>1000
Melting point (°C)	1614
Dielectric constant	7.40–10.47
Thermal conductivity ($W/cm \cdot K$)	0.013
Chemical composition	$Ca_{10}(PO_4)_6(OH)_2$
Ca/P molar	1.67
Color	White

ions within the solid. The second method is the wet-chemical method, which is relatively easy to conduct, even at low temperature, in contrast to the solid-state reaction. As per publications, this is one of the simplest methods frequently used to prepare HA. The materials prepared by this method are usually poorly crystallized and homogeneous in composition. Hydrothermal production methodology involves the reaction in aqueous solution of precursors at high temperature and high-pressure atmosphere. This method enables the preparation of a well-crystallized and homogeneous phase. The product obtained by this method is easily sinterable due to the effects of high-temperature high pressure aqueous solutions. More recently, ceramic powders are being prepared using microwave techniques under optimum conditions in order to minimize the reaction time, production cost, and to increase the purity of the sample [212].

One of the potential properties, namely, high surface area reactivity, associated with nanocrystals can be exploited to improve many biomedical applications [154, 51, 52, 57]. In addition, nanocrystallized characteristics have proven to be of superior biological efficiency [219]. For example, compared to conventionally crystallized HA, nanocrystallized HA promotes osteoblast cell adhesion, differentiation and proliferation, osteointegration, and the deposition of calcium-containing minerals on its surface [218], and also enhances the formation of new bone in a short period. Therefore, nanostructured HA is thought to represent a unique class of biomaterial formulations that promises enhanced implant bonding and improved overall implant efficacy. On this topic, we discuss the recent advances made on nanoscale bioactive HA. Recent publications concerning HA showed that there are few possible methods available for the preparation of nanoscale HA (Table 5). Toriyama et al. succeeded in obtaining HA by starting from brushite and calcium carbonate [207], and then Yokayama et al. [230] doped some trace elements such as Sr to HA in order to strengthen it mechanically. They obtained HA from mechanochemical treatment with low crystallinity. Further, they postheated the samples at temperatures between 500 and 1000 °C in order to promote their crystallinity. Free-standing crystalline nanostructured HA was synthesized using a wet-chemical technique [53]. They consolidated powder agglomerates into bulk form for the testing of fracture toughness, bending, and compressive strength. Their results clearly showed that the compressive strength of bulk nano-HA could be increased up to 4 fourfold in comparison to bone (879 versus 193 MPa for compact bone), whereas the bending strength was found to be much equivalent (913 versus 160 MPa in bone).

A novel method for the preparation of HA fibers was introduced using a solid phase reaction [168]. The HA fibers were fabricated by heating a compact consisting of calcium metaphosphate fibers with calcium hydroxide particles at 1000 °C in air atmosphere, and subsequently treating the resultant compact material with dilute aqueous HCl solution to remove unwanted secondary phase substitution such as

Table 5. Various methods for nanohydroxyapatite preparation.

Techniques	Grain size	General description	Ref.
Solid state	~0.5μm	Inhomogeneous, large grain size, irregular shapes, reaction condition 900–1300 °C	[168]
Wet chemical	20–200 nm	Micro- to nanograin size, low crystallinity, homogeneous, reaction condition: room temp. to 100 °C	[53, 47, 232, 170, 121]
Precipitation/ hydrothermal	11–24 nm	Homogeneous, micro/nanosize, low crystallinity, reaction condition: room temp. to 200 °C (1–2 MPa)	[234]
Hydrothermal	60–70 nm	Homogeneous, ultrafine crystals, high temperature, and high-pressure atmosphere	[107, 85]
Mechanochemical	<20 nm	Easy production, low crystallinity, ultrafine crystals, room temp. process	[230, 161]
pH shock wave	20–100 nm	High-energy dispersing, nonporous, monocrystalline particles with Ca/P molar ratio 1.43–1.66	[122]
Microwave	115–300 nm	Time and energy saving, uniformity, fine grain size	[212, 229]

CaO by acid leaching. The obtained HA fibers were characterized systematically, and the morphological structure was evaluated by SEM analysis. The structural morphology of HA fibers clearly showed nanostructural characteristic features more or less similar to natural bone mineral apatite phase. Zhang and Consalves [234] reported the thermal stability of nano-HA prepared by the precipitation method, followed by hydrothermal treatment. By this approach, the obtained HA has a rod-like shape, with an average crystal size of 11–24 nm with respect to the precursor's compositions. They produced various calcium phosphates of HA powder with respect to the starting material from 1.66 to 1.73 (Ca/P ratio). The HA also was prepared in the form of lath-like monocrystalline particles using high-energy dispersing equipment in combination with a pH shock-wake method using $Ca(H_2PO_4)_2{\cdot}H_2O$ and $CaCl_2$ with a Ca/P molar ratio of 1.66, almost similar to the stoichiometric Ca/P ratio of HA, and characterized using various analytical methods to determine their phase composition, structure, and texture [122]. They measured the Ca/P molar ratio from 1.43 to 1.66 with respect to the method of preparation of HA, and found that the pH shock-wave method produces a nonporous monocrystalline HA lath with particles with an average grain size of ~140–1300 nm in length, 20–100 nm in width, and 10–40 nm in thickness. The obtained Ca/P atomic ratio is 1.66, which is very close to the theoretical value of stoichiometric HA (1.67). Per their investigation, it seems to be a unique method since it creates the proper hydrodynamic conditions for lath-like particle growth in the [001] direction with the physicochemical characteristics that could not be obtained by using conventional mechanical stirring equipment. Ahn et al. [4] derived a bioactive nanocrystalline HA through nanostructure processing to achieve superior chemical homogeneity and microstructural uniformity in order to get high-quality apatite-based bioceramics. They attained nanostructural HA by the precipitation method, and studied the optimal conditions by changing the reaction parameters such as precursor pH, aging time, and temperature to achieve nano-HA particles with tailored composition and morphology. This investigation clearly showed that the nanostructure processing offers an alternative solution to challenges encountered in the HA processing; they also concluded that the nanostructured HA results in a 70% improvement over the best conventional HA. In 2002, Driessens et al. [47] reported a new class of nanoapatite cement having Ca/P ratio in the range of $0.8 < Ca/P \leq 1.5$, which is entirely different from the conventionally prepared calcium phosphate cement ($1.5 \leq Ca/P < 1.8$). They found that the prepared nanoapatitic calcium phosphate cement first will be transformed into CO_3–Ca deficient HA before it can be resorbed in the bone. After that, it will only be transformed into bone as a function of commercial calcium phosphate cements. However, there is no detailed report of animal experimentation on this product. The preparation and characterization of nanosized HA under microwave irradiation were investigated recently [212]. The nano-HA was synthesized by the precipitation method using $NH_4H_2(PO_4)$ and $Ca(NO_3)_2{\cdot}4H_2O$. The obtained precipitate was subjected to microwave sintering after washing with distilled water using a high-speed centrifuge. The microwave generated a power conception of 980 W and a power output of 490 W with frequency 2.45 GHz for a period of 5–15 min. This method offers advantages over normal furnace heating with respect to time and energy saving. They found that there was a slight increase in the microhardness value of the sintered material with an increase in microwave exposure time. The major problem associated with the microwave densification is in getting large samples without microcracks. This is believed to be due to the nonuniform packing of crystallite in the green compact, as well as due to the nonuniform heating profile occurring inside the microwave cavity.

After the development of nanoscale HA, its application to coating on bioimplants such as titanium alloy has been increased. The coating, conventionally and commercially, so far has been done by the thermal-spraying method, which functions as an intermediate layer between human tissues and the metal implant. This type of coating method can speed up early stages of the healing process after postimplantation, but the lifespan is very limited due to the low interfacial bond strength between the tissue and the implant, which arises from the dissolution of amorphous HA in human body fluid. To alleviate these problems associated with interfacial strength, Zang et al. [232] developed a novel room-temperature electrophoretic deposition process to fabricate nanostructural HA coating. This kind of nanostructured HA coating significantly improved the coating's bond strength up to 50–60 MPa, that is, two–three times better than the thermal-sprayed HA coating, and has excellent corrosion resistance, 50–100 times higher than the conventional HA coating. Apart from this, nanostructured carbonated HA also has been used as a coating material on titanium alloy using the plasma-spraying technique [85]. The acquired results showed that the sprayed precursors consisted of β–$Ca_2P_2O_7$, α–$Ca_3(PO_4)_2$, and $CaHPO_4$ phases, whereas the carbonated HA coatings exhibited a unique phase construction of nanostructural and needle-like crystals. They found that the dissolution and bonding strength degradation of the carbonated HA coatings were much lower than the plasma-sprayed HA coatings. The structural characterization of the carbonated HA coating obtained in this study is expected to show much more biological efficacy.

2.2. Polymeric Nanomaterials

The design of polymeric structures at the nanometer-scale level has received great attention because their submicron size particles could deliver biomolecules in a controlled fashion. The use of polymers in medical applications dates back almost to the birth of polymer science; since then, they have been used for the design of medical prostheses or implants and for the delivery of therapeutic molecules [14, 105]. The interest in polymers for various biomedical applications is mainly due to their design flexibility (because the polymer composition and structure can be tailor made to a specific need) and their biodegradation in body pH. Polymers can include chemical bonds that undergo hydrolysis upon exposure to the body's aqueous environment, and they can also degrade by cellular or enzymatic pathways. The rate of degradation can be controlled by manipulating the polymer properties such as hydrophobic and crystalline. Polymers can be broadly

classified into two groups, depending upon their degradation behavior, namely, biodegradable and nonbiodegradable polymers (Table 6). Considering the mechanical properties of the polymers (Table 7) with the host tissue, most of the polymers frequently have been used for soft-tissue applications [180]. Polymeric systems that have been investigated for hard-tissue regeneration include poly(hydroxy esters) [146–148, 204, 205, 202], polydioxanone (PD) [162], poly(propylene fumarate) (PPF) [72], poly(ethylene glycol) (PEG) [148], poly(orthoesters) [197], polyanhydrides [142], and polyurethanes (PU) [163]. Among the poly(hydroxy esters), PLA, PGA, and PLGA serve as scaffolds that have an appropriate three-dimensional (3D) structure on which osteoblast cells can grow and deposit bone extracellular matrix (osteoid). The polymers degrade after the osteoid has become established. These polymeric materials may also serve as carrier vehicles for the delivery of bioactive molecules, and at the same time, they have a structural function. In addition, poly(L-lactic acid) has the highest mechanical strength as compared to PGA and PLGA, and the same was investigated for orthopedic applications in the form of pins, screws, and plates. The nondegradable polymers also were investigated in many hard-tissue applications, including poly(methylmethacrylate) (PMMA) [46], poly(hydroxyethylmethacrylate) (PHEMA) [79], poly(acrylic acid) (PAA) [114], poly(butylmethacrylate) (PBMA) [132], and poly(glycidylmethacrylate) (PGMA) [156–159].

For the past few years, there has been growing interest in a novel polymer fiber production methodology. Polymeric fibers having a high degree of orientation along the fiber axis are fabricated from polymer melt or solution through a variety of processes, including drawing or spinning by involving a number of complex molecular processes [40, 8, 191]. These processes involve time- and temperature-dependent molecular motions, phase transitions under highstress, entanglement constraints, and various intermolecular reactions. Hence, the final state of molecular order in a fiber is a function of the process variables such as stress, strain, temperature and time, and the length and length distribution of the molecules. The polymer fibers developed by these techniques have a diameter in the range of 1–100 μm. Recent advances in the fabrication of nanofibers offer another form of building block for the nanostructure-based biomaterials [138]. Polymer nanofibers can be produced by a number of techniques such as drawing [165], template synthesis [139, 61], phase separation [136, 228], self-assembly [130, 222], and electrospinning [75, 63, 96]. Drawing is a simple method that is being explored to prepare polymeric nanofibers because it needs minimum equipment, and is also feasible for producing lengthy fibers. By this process, it is possible to create nanofibers with diameters in the range between 2 and 100 nm approximately. The only disadvantage of this process is that it produces fibers one after another. Template synthesis is also a considerably easy method to produce nanofibers. This method uses a nanoporous membrane as a template during the fiber formation. The highlight of this method is the fabrication of various sizes of solid as well as hollow-shape fibers by optimizing the experimental conditions. However, this method is limited to a few polymers only, which may be disadvantage of the template synthesis. Phase separation is an interesting method involving various processing steps, including material dissolution, gelation, solvent extraction, freezing, and drying. Of course, to carry out these steps would take a prolonged time to convert the polymer into a nano form. The advantage of this method lies in the controlling of mechanical properties of polymer nanofibers. Self-assembly is a process in which individual, preexisting components organize themselves into desired patterns and functions. It is also a time-consuming process. Ultimately, the legend of the polymer nanofiber processing technique may be termed so-called electrospinning.

Table 6. Types of polymers and their applications with respect to degradation behavior.

Materials	Applications
Nonbiodegradable polymers	
Polymethylmethacrylate (PMMA)	Bone and dental cement, intraocular lens
Polyhydroxyethylmethacrylate (PHEMA)	Drug delivery
Polybutylmethacrylate (PBMA)	Bone and dental filling, controlled release
Polyacrylic acid (PAA)	Bone plates, tissue engineering
Polypropylene (PP)	Heart valves, skin patches, sutures
Polytetrafluroethylene (PTFE)	Vascular grafts, reconstructive surgery
Polyamides	Drug delivery, catheters, syringe
Polycarbonate (PC)	Contact less, syringe, skull implants
Biodegradable polymers	
Polylactic acid (PLA)	Sutures, bone plates, screws, scaffolds
Polyglycolic acid (PGA)	Sutures, bone plates, screws, scaffolds
Polylactic-co-glycolic acid (PLGA)	Tissue engineering, drug delivery, cartilage screws, sutures, wound dressing
Polycaprolactone (PCL)	Drug delivery, bone plates, fixation device
Polyhydroxybutyrates (PHB)	Drug delivery, bone plates
Polydioxanone (PD)	Sutures
Collagen	Prosthetic heart valves, membranes, wound dressing, tissue engineering
Gelatin	Binding agent, tissue engineering, membranes

Table 7. Mechanical properties of polymeric biomaterials [59, 13].

Polymers	Tensile strength (MPa)	Tensile modulus (GPa)	Flexural modulus (GPa)	Glass transition (°C)	Melting temperature (°C)
Poly(L-lactic acid) (Mw:100,000)	50	2.7	3.0	58	169
Poly(D,L-lactic acid) (Mw:107,000)	29	1.9	1.95	51	n/a[a]
Poly(glycolic acid) (Mw:50,000)	n/a	n/a	n/a	35	210
Poly(caprolactone) (Mw:44,000)	16	0.4	0.5	−62	57
Poly(β-hydroxybutyrate) (Mw: 422,000)	36	2.5	2.85	01	171

[a]n/a = Not available.

For the past few years, there has been growing interest in novel fiber production by electrospinning. It is a unique process because it can consistently generate nanoscale fibers ranging between 50 and 500 nm in diameter [43]. The electrospinning can be considered a variation of the better known electrospray process. The electrostatic-spray techniques enable the understanding of the insights of the electrospinning process, which involves applying a high voltage to a capillary and pumping a polymer solution through it. As a result, nanofibers of a polymer collect as a nonwoven fabric on a grounded plate below the capillary. The driving force of the mechanism behind the spinning of nanofibers is when, in the absence of the electric field, the fluid forms a drop at the exit of the capillary, and its size is determined by surface tension. If an electric field is applied, it stimulates charges into the solution, and quickly relaxes to the fluid surface. The coupling of the surface charge and the external electric field creates a tangential stress, resulting in the deformation of the droplet into a conical shape, better known as a Taylor cone. Once the electric field exceeds a critical value needed to overcome the surface tension, the apex of the cone ejects a fluid jet. A schematic representation of the electrospinning process is shown in Figure 3. This process has three stages. First, there is a jet initiation and extension along a straight line. The second step pertains to growth of a bending instability and elongation of the jet, which allows the jet to become very long and thin while it follows a looping and spiraling path. This is followed by, of course, the final step, solidification of the jet into nanofibers. The background development of this electrospinning process recalls many early inventions as when Lord Rayleigh [134] studied the instability that occurred in electrically charged liquid droplets. He determined the maximum amount of charge of liquid droplets before the applied electric forces overcome surface tension of the droplets, followed by Zeleng [233], who described the phenomenon of the electrospraying technique. Later, Taylor [233] analyzed the conditions at the print of a droplet that is deformed by an applied electric field, and reported that a conical interface between two fluids is stable if the cone has a semiangle of 49.3°. These findings have a great impact on the process of electrospinning of a polymer solution or melt. However, the electrospinning of the solution of macromolecules can be traced back to 1934, when Formhals invented a process for making polymer fibers by using electrostatic force. He found that the electrostatic potential difference required depended upon the properties of spinning solution, such as molecular weight and viscosity [64–70].

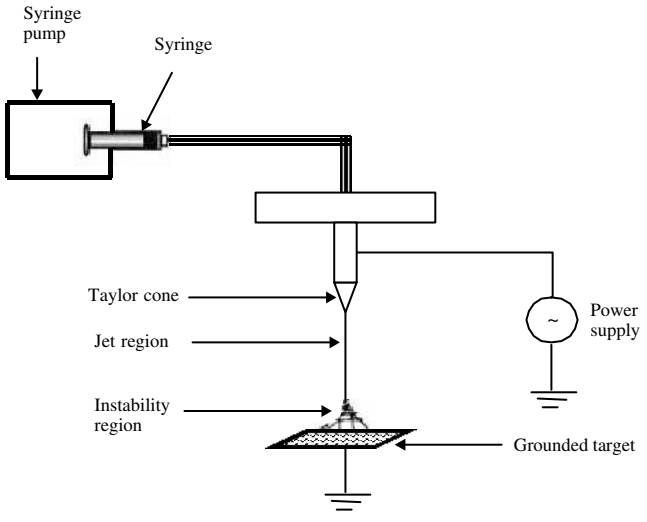


Figure 3. Schematic representation of electrospinning process.

Because of the small pore size and high surface area, electrospun nanofibers have shown promise for exploitation in protective clothing, molecular filtering [43, 183, 75, 74], and in many biomedical applications [21, 75]. In addition, polymer nanofibers are being explored in making optical and electronic devices [164, 26], and as reinforcement agents in nanocomposites [26]. The polymeric nanofibers formed by the electrospinning technique have been proposed for soft- and hard-tissue applications, depending upon their structural and mechanical properties. Also, many investigations suggested the use of polymer nanofibers as medical prostheses, including blood vessels, vascular, breast treatments [140, 141, 95, 16, 11, 189, 198], and for hard-tissue remodeling [6, 21–23]. Polyethylene oxide (PEO) was fabricated in the form of nanofibers by electrospinning for wound-healing applications [38]. Dashy and Reneker studied the processing conditions and fiber morphology, and found that the diameter of the fibers could be varied by changing the processing parameters such as electric field, distance between capillary and the collection plate, and viscosity of the polymer solution. This kind of nanofiber could

be used as a wound-dressing material and as a reinforcement agent in making nanocomposites. Hulteen et al. [98] reported the preparation method of carbon nanotubes and nanofibrils with a nominal outer diameter of 200 nm and a length of 60 μm. They characterized their products with SEM, and found that the structure was extended throughout the entire length of the membranes (~60 nm) without any sign of breakage. Carbon nanotubes can also be catalytically produced in large quantities, and have been used for the reinforcement of nanostructural composite materials [175]. Elias et al. [55] enhanced the functions of osteoblast cells on carbon nanofibers. They investigated the *in vitro* performance of osteoblasts (the bone-forming cells) onto carbon nanofibers, and found that the osteoblastic proliferation increased with decreasing carbon fiber diameters after a few days of culture. The comparative study of the carbon nanofibers prepared by electrospinning with respect to conventionally prepared carbon fibers with dimensions larger than 100 nm implies that the nanoscale fibers represent a unique and promising class of material for the cell culture of osteoblastic cells. These results explored the ability of carbon nanofibers to improve the osteointegrative properties in a controlled manner. The effect of bone-cell adhesion on novel carbon nanofibers and polycarbonate urethane/carbon nanofiber composites was investigated at the *in vitro* level [178]. The results provided the evidence that the smaller scale (nano) carbon fibers promoted the osteoblast adhesion. The adhesion of other cells was not influenced by the carbon fiber dimensions. Also, smooth muscle cells, fibroblast, and chondrocyte adhesion decreased with an increase in either the carbon nanofiber surface energy or the simultaneous change in the carbon nanofiber chemistry. The versatility of these fibers suggested that there are a large number of possibilities for future designs that could enhance the efficacy of medical implants.

3. NANOMATERIALS IN TISSUE ENGINEERING

The field of tissue engineering has emerged over the past decade by evolution of the engineering discipline in combination with materials and molecular cell biology. Tissue engineering may be defined as the application of engineering disciplines to either preserve existing tissue structures or to allow tissue growth. Although tissue engineering has multiple facets and several definitions proposed so far, it is also known as the science of persuading the body to regenerate or repair tissues that fail to regenerate or heal spontaneously [2]. Tissues are usually organized into three-dimensional structures in our body. This architecture is increasingly believed to contribute considerably to the growth of specific biological function in tissues, and it is thought to provide appropriate nutritional conditions and a spatial environment for cells [12]. All of the materials intended for human applications undergo tissue responses upon implantation into living tissue with respect to the implanted site. Recently, progress has been made in the design of nanostructural biomaterials capable of interacting very closely with the biological environment. This interaction has been accomplished by harnessing the principles involved in biological recognition between proteins and proteoglycans on a cell surface with other biological molecules normally present in the natural extracellular matrix (ECM). An ECM is a complex matrix of proteins, proteoglycans, and charged polysaccharides, which provides physical scaffolding for cells and tissues, and further helps to provide a permeability barrier between tissue compartments, and hence enables the polarization of tissue structures [82]. The ECM molecules are typically quite large, of molecular weight above 100,000, and contain several functional domains or regions, including those involved in receptor-mediated cell adhesion and regions which interact with other ECM molecules or growth factors. These biomolecules have a notable role in tissue repair or regeneration by allowing cell-free materials to elicit cellular infiltration within implanted materials, and to direct the formation of new tissues within the material at the implanted site. By this technique, the implanted biomaterial functions as a mechanical supporter, which is able to alter the track of tissue regeneration through the transmission of biological signals to cells in the surrounding tissue.

Generally, tissue can be classified into hard and soft, depending upon its mechanical and structural characteristic features. For example, bones and teeth pertain to a hard kind of tissue, which is stiffer (elastic modules) and stronger (tensile strength) than the soft tissue. Skin, blood vessels, cartilage tendon, and ligaments are a few notable examples of the latter type. Table 8 shows the major differences between hard and soft tissues in terms of mechanical properties [180]. All hard and soft tissues of our body contain differentiated cells living in an ECM elegantly designed by nature's master mind. These matrices consist of nanofibers, microfibers, sheaths of material, minerals, proteins, polysaccharides, and other essential biopolymers. The nanoscale architecture of these matrices is critical for the proper function of each tissue. It is reported that the cells can attach and organize well around fibrous scaffolds, even though the fiber diameter is smaller than the cell [124]. Hence, it is clear that the nanoscale fibrous matrix as an artificial scaffold could provide better circumstances for the growth of cells in a natural way. Progress in nanoscience and nanotechnology is expected to enable the fabrication of structurally and functionally designed artificial nanoscaffolds that will support cells to live in a natural environment with all

Table 8. List of mechanical properties of hard and soft tissues [180].

Tissues	Young's modulus (MPa)	Tensile strength (MPa)
Hard tissues		
Cortical bone	7,000–30,000	50–150
Cancellous bone	50–500	7.4
Dentine	11,000–17,000	21–53
Enamel	84,000–131,000	10
Soft tissues		
Articular cartilage	10.5	27.5
Fibrocartilage	159.1	10.4
Ligament	303.0	29.5
Tendon	401.5	46.5
Skin	0.1–0.2	7.6

of the necessary qualities. In this regard, a fundamental understanding of cell–substrate interactions is an important criterion for tissue remodeling in the development of novel biomedical implants. It has been hypothesized that cell–substrate interactions are governed by surface properties and their imperative interactions found to occur within 1 nm of the implant surface [117]. Naturally occurring biological structures appear to interact selectively with relevant biomolecules while resisting nonspecific interactions. As biological interfaces are highly dynamic, simply trapping cells at a particular point on a surface is not enough [15]. Cells must first be encouraged to differentiate, that is, change their behavior to perform as required, and once their function is complete, their activity must be turned off again. Transient interactions are thus to be encouraged, and multiple-function surfaces are likely to prove extremely important. To gain insight into these interactions, a logical approach is needed to investigate the scaffolds on which the cells attach and grow in living systems. To make this happen, material scientists, biologists, engineers, and clinicians will have to work together to create the necessary strategies to synthesize scaffolds at the nanolevel.

The scaffolds or templates can be fabricated through the use of lithographically defined patterning or processing of the substrate materials in order to achieve selective nucleation and growth of the nanostructured biomaterials similar to natural ECM. Hanarp et al. [86] studied the adsorption of 110 nm negatively charged polystyrene particles on positively charged titanium surfaces by colloidal lithography for the purpose of making surfaces with controlled nanotopography for the cells' adhering. The coverage and structure of the particle films were in good qualitative agreement with the soft sphere random sequential adsorption model. They found that the colloidal lithography is a quick and relatively simple method to produce uniform nanostructures over large surface areas. However, there is also a host of natural templating materials that researchers can avail themselves of. Recently, a biodegradable polymer nanofiber scaffold was prepared by the electrospinning process using poly(caprolactone) (PCL) for bone tissue engineering [231]. In this process, PCL fibers with diameters down to the nanometer-scale rage were formed by subjecting a fluid jet to a high electric field. The biocompatibility of these nanofibers was evaluated with mesenchymal stem cells derived from the bone marrow of neonatal rats. The cell–polymer constructs were cultured with osteogenic supplements under dynamic culture conditions. The potentiality of PCL nanofibers for tissue engineering was confirmed by the penetration of cells into the nanofibrous scaffold. This kind of scaffold provides an environment that supports mineralized tissue formations, and may also be used in the treatment of bone defects. Apart from this, other possible micro/nanostructure scaffold fabrication techniques are reviewed in a detailed manner [135, 203] as illustrated in Table 9. Carbon-based nanosystems and carbon nanotubes are another class of materials which provided a tremendously rich basis for nanostructure research. Carbon fibers are a new breed of superstructure materials mainly used in composites as reinforcing agent. Carbon fibers offer the highest specific modulus and strength of all reinforcing fibers thus far used in nanocomposites [28]. They can also serve as a precursor in the fabrication of nanomaterials such as starting material or can provide an avenue for the synthesis of natural templates of varying sizes depending upon medical applications. This kind of scaffold can also be seeded with appropriate cells, growth factors, or therapeutic macromolecules to enhance their properties, and hence widen their usage. Carbon fibers have wide medical applications, either as such or as composites, and have been investigated extensively. Their potentiality as molecular filters for the separation of biological molecules is also noteworthy. Recently, carbon fibers were produced by the electrospinning technique at the nanolevel. Wang et al. [216] prepared ultra fine fibers from a polyacrylonitrile/N,N–dimethylformamide precursor solution using electrospinning. They made carbon fibers on the order of 100 nm diameter. This study particularly emphasized the chemical analysis features of carbon nanofibers by Raman spectroscopy. This kind of carbon nanofibers could be used as cell carriers in tissue engineering applications.

Table 9. Various fabrication methods for tissue scaffolds.

Methods	Ref.
Solution casting	[195]
Gel casting	[34]
Solvent casting and particulate leaching	[146]
Vibrating particle and particulate leaching	[3]
High-pressure gas foaming and particulate leaching	[89]
Emulsion freeze drying	[214, 215]
Membrane lamination	[223]
Gas saturation	[153]
Lyophilization	[101]
Three-dimensional printing	[76]
Fiber bonding	[147]
Phase separation	[136]
Template synthesis	[61]
Drawing	[166]
Self-assembly	[222]
Electrospinning	[231]

As compared to other types of biomaterials, biodegradable polymers are frequently being used in tissue engineering due to easy fabrication and their potentiality in cell attachment. The list of nanostructure biomaterials used as scaffolds in tissue engineering applications is given in Table 10 with respect to the types of cells studied. The essential properties of an ideal scaffold are that it should be biocompatible, biodegradable in a controlled fashion, have an appropriate pore size with interconnective porosity to achieve better cell growth, be surface modifiable, have adequate mechanical strength, the potential to elicit an immunological or foreign body reaction, should encourage the formation of ECM by promoting cellular functions, and should possess the ability to carry biomolecular signals such as growth factors. Scaffolds intended for tissue remodeling must direct the arrangement of cells in an appropriate three-dimensional configuration, and present molecular signals in a suitable spatial and temporal manner so that the individual cells will form the desired tissue structures. This should be easily fabricable, and reproducible within economical limits. Most of the degradable polymers are at least cell adhesive [25, 35, 171, 211, 210, 54]. The scaffolds

Table 10. List of nanostructured biomaterials used as tissue scaffolds with respect to cell type.

Scaffold materials	Grain size	Cells studied	Ref.
Carbon fibers	<100 nm	Osteoblast	[55]
Polycarbonate urethane/carbon fibers	<100 nm	Osteoblast	[178]
Collagen/PEO	—	Intended as T.S[a].	[96]
Collagen	—	Intended as T.S.	[144]
Collagen–chitosan	Nanoscale	K562, hemopoietic cell	[199]
Hydroxyapatite	115 nm	Osteoblast-like cells	[229]
Hydroxyapatite	<100 nm	Mesenchymal stem cells	[137]
Hydroxyapatite	160–220 nm	Intended as T.S.	[133]
Hydroxyapatite	—	Intended as T.S.	[107]
Hydroxyapatite	100–200 nm	Cell carrier	[121]
Hydroxyapatite	10–500 nm	Angiogenic factors	[73]
Titanium-coated silicon/PMMA	0.2 μm	Cerebral neuron	[32]
Alumina/TiO_2	23–32 nm	Osteoblast	[217]
Silk/PEO	—	Intended as T.S.	[112]
PLA/PGA	—	Intended as T.S.	[129]
PLGA	—	Osteoblast	[201]
PLA	~200 nm	Nerve stem cells	[228]
PCL	20 nm–5 μm	Mesenchymal stem cells	[231]
Polystyrene/silicon	0.5 μm	Calavarial cells	[29]
Polystyrene/epoxy	2.0 μm	Fibroblasts, epithelial	[165]
Polycarbonate	0.5–2.5 μm	Fibroblast	[100]
Polyurethane/PMMA	Nanoscale	Endothelial cells	[77]
Polyurethane	~100 nm	Endothelial cells	[77]

[a]T.S. = tissue scaffold.

are frequently made by using biodegradable polymers, either natural such as collagen, fibronectin, elastin, gelatin, or synthetic such as PLA, PGA, PLGA, and PCL. Some of the composite materials are also available in conjunction with HA for the purpose of bone tissue engineering. As the function of a nanofibrous scaffold is that it mimics the natural ECM, the conversion of biomaterials (in particular, polymers) as nanofibers has been enhanced rather than conventional larger diameter fibers (few microns), which do not have the morphological structure of natural tissue. Huang et al. [96] reported the high-resolution analysis of engineered type I collagen nanofibers generated by the electrospinning process. As collagen is a primary structural element in the ECM, they intended this nanofiber for tissue engineering applications. Collagen nanofibers were generated at ambient temperature and pressure by electrospinning a 1 wt% solution of type I collagen and PEO, and the fiber features were studied by a scanning electron microscope (SEM) and a transmission electron microscope (TEM). They also renewed emphasis on the use of processing of natural materials, which is the one they addressed in this experiment. This strategy may lead to the production of collagen scaffolds for the repair of connective tissue injuries, and as a matrix for cellular growth directed at organ regeneration. As collagen has a variety of tissue remolding applications, it is necessary to spin those natural polymers into nanofibers in order to widen their usage. To address this, Matthews et al. [144] fabricated a collagen nanofiber matrix using types I and III. They demonstrated the possibilities of tailor-made subtle mechanical properties into a matrix by controlling the fiber orientation. The inherent properties of the electrospinning process make it possible to fabricate complex and seamless three-dimensional structures. They proved that those natural polymers promoted cell growth, and achieved considerable cell penetration into the engineered collagen matrix. They suggested that this kind of nanofibrous matrix may represent a considerable tissue engineering scaffold due to their structural and biological features.

In the role of bioceramics, nano-HA has been used in tissue engineering applications, and a few investigations also reported its activity with cells. Yang and Ong [229] prepared nano-HA by the microwave processing technique, and cultured osteoblast-like cells on the surface of HA. They used a 115 nm particle size of HA for this *in vitro* evaluation, and observed the osteoblast cell morphology behavior. There were no significant differences of cell proliferation on the HA derived from conventional as well as microwave processes, and no sign of apparent impairment of cell physiology also was detectable. These results showed that the microwave processing condition did not have any apparent detrimental effect on the biological properties of HA. The *in vitro* response of sol–gel-derived HA with human mesenchymal stem cells was described in order to prove their enhancement of bioactivity and biocompatibility [137]. The *in vitro* results showed that the human mesenchymal stem cells adhere to aero–sol-derived HA coatings, and exhibited better proliferation signals on these surfaces. The evidence of the proliferation rate of mesenchymal stem cells was observed with a tendency to spread onto the surface as if forming groups, thus supporting the notion of not only biocompatibility, but also enhanced bioactivity of the materials at the *in vitro* level. This behavior is considered to be of great importance since it simulates in an artificial surface several steps of the cell cycle occurring in a tissue environment with no need for an interfacial protein scaffold. They concluded that these surfaces present an interaction with the mesenchymal stem cells that may be useful for tissue engineers to design an alternate to biological ECMs. A new class of nano-HA in the form of fibers is also being used in many biomedical applications, in particular, tissue engineering. The production of such a type of nanofiber was reported recently by a simple route using the reaction between calcium chloride and sodium phosphate in the presence of a surfactant [133]. Liu et al. produced a high-crystalline and fairly pure phase of HA with length and diameters of 160–220 nm and 5 nm, respectively. One more interesting publication is one in which Ioku et al. [107] reported the preparation of fibrous apatite and apatite sheet by the hydrothermal process intended for both bone filling and tissue engineering applications. Since HA is a bioactive agent, it will act both as a cell carrier and as an osteoconductive substance at the defective part of the bone tissue engineering. This kind of HA sheet could be used as a matrix for bone tissue engineering, and for the delivery of bioactive molecules due to its biocompatibility and biostability. A comprehensive article by Kizuki et al. [121] demonstrated the effect of polarization on cell behavior by cell cultivation on electrically polarized ceramic HA. The acceleration of the cell proliferation was confirmed on the surface of the polarized ceramic HA. The cell proliferation was accelerated

by the bone-like layer on the material. It was found that the bone-like layer significantly reinforced the cell adherence, and the cells on the surface of polarized HA were more proliferated than those on the ceramic HA. Porous biomaterials are of particular interest in tissue engineering as cells/tissues invade the pores for their growth upon implantation. Gibson et al. [73] addressed the porous structure biomaterials by producing an interconnected membrane-like web of small fibers with a diameter size in the range between 10 and 550 nm suitable for tissue engineering applications. Tissue reaction to the implant is determined by the combined effects of the implant material, structure, and carrier fluid. Induction or enhancement of biological reactions are achievable by bioactive implant surfaces or by bioactive substances released from the implant materials or included in the carrier fluid. For instance, the inclusion of angiogenic factors within the matrix may serve to enhance its vascularization. To induce bone in-growth into the scaffold, a bioactive agent, that is, a calcium phosphate kind of nanoscale material, can be included. Recently, research has focused on the incorporation of inorganic bone fillers such as hydroxyapatite [190] and tricalcium phosphates [102] into a polymeric matrix to produce bioactive tissue engineering materials. These inorganic fillers allow tailor-making the desired degradation and resorbtion kinetics of the polymer matrix. More importantly, the basic resorbtion products of HA or TCP would buffer the acidic resorbtion byproducts of the polymers (e.g., aliphatic polyesters), and may thereby help to avoid the formation of an unfavorable environment for the cells due to decreased pH [1]. These kinds of biomaterials improved the biocompatibility and tissue integrity upon implantation with surrounding tissues.

4. NANOMATERIALS IN DRUG DELIVERY

Over the past few years, there has been a tremendous advancement in the development of drug delivery systems. Recently, many efforts have been taken into consideration to design nanoscale biomaterials as drug carriers. It may consist of solid colloidal particles ranging from 10 to 1000 nm grain size. Nanotechnology is currently developing as an advanced conceptual approach in medical and pharmaceutical applications due to its flexibility of materials design to deliver therapeutic macromolecules or other bioactive molecules. Nanoparticles have been found to be of much use for the controlled release of therapeutic molecules such as drugs and other bioactive molecules, including growth factors [84, 103]. As nanoparticles have a large surface area, it is expected that they will release the therapeutic macromolecules or growth factors in a controlled fashion, in particular, target delivery systems. In this area, a major concern is that, when delivered, the molecules in physiological fluids such as blood, or organs like lungs, interact more with the surroundings of the body environment. As such, the surface characteristics of the particles must be properly designed to maximize the potential for favorable interactions, and their preparation procedure must be delineated accordingly. In the area of medical applications, finely dispersed pharmaceuticals offer rapid drug delivery and reduced dosages for patients [177]. Expensive health-care materials such as drugs may be efficiently dispersed in small and controlled quantities through the nanoarchitecture of carrier materials [188]. The preparation process of carrier materials is generally achieved by the liquid dispersion method. The prepared materials in this approach are being widely used to apply topical coatings to the human epidermis because they can be absorbed faster and more completely than the conventional coating method. Recently, it was found that the HA nanoparticles have good affinity in the delivery of therapeutic macromolecules and growth factors. In this direction, porous HA is used for the intestinal delivery of insulin [173]. Insulin was loaded into nanoparticles by soaking HA in a fixed amount of insulin at 2–6 °C and encapsulated with a sodium alginate solution. The insulin-loaded HA nanoparticles encapsulated in solution alginate can efficiently release the insulin in the neutral environment of simulated intestinal fluid. Hence, the delivery system with HA nanoparticles encapsulated with an alginate matrix seems to be a suitable carrier for the oral delivery of insulin.

The delivery of therapeutic macromolecules in a controlled fashion, either as a sustained release or a targeted release to a patient's body, is of great importance in drug delivery systems. Some of the researchers identified that the particle size and shape greatly influence the delivery of macromolecules. It is appreciable if the particle size is less than 5 μm to circulate through the smallest capillaries, and a particle diameter less than 200 nm is needed to avoid spleen filtering effects [149, 81]. The particulate drug carriers tend to be readily cleared, and accumulate upon delivery at the tissue sites, in particular, liver and spleen. For example, polystyrene (PS) [104], albumin [187], PLA [10], and PLGA [181, 80] are frequently used as drug carriers. These are all mainly used for target delivery systems. The potential applications of PEG-coated nanoparticles as effective drug carriers also have been studied in a detailed fashion by many researchers [169, 182, 58, 7]. Recent publications showed that a few drugs have been encapsulated in PEG-coated nanoparticles, and their release kinetics were studied. Lidocaine was successfully encapsulated as a model drug to study the possibility of controlling the rate of drug release from nanospheres of PEG–PLA, PEG–PCL, and PEG–PLGA copolymers [175]. The results obtained indicated that the drug release was governed by the nature of the core and by the coating layer of polymer, particularly PEG. Gorner et al. [78] loaded lidocaine in PLA and PEG–PLA nanospheres, and studied the drug release kinetics. They found that the drug release was directly related to the size of the nanoparticles and the dispersion of the drug inside the particles. Matsumoto et al. [143] reported the loading of progesterone into PLA and PLA–PEG–PLA copolymeric systems, and studied the rate of drug loading and release. They found that the release kinetics and polymer degradations significantly pertained to the PEG content in the copolymeric carrier. It was also emphasized that the polymer dispersity plays a notable role in the drug release mechanism. Recently, there has been a growing interest in the delivery of drugs or growth factors using polymer nanofibers. The significant role behind the delivery of the biomolecules is based on the principle that the rate of dissolution of the drug increases with the increasing surface area of the drug as well

as the carrier. As nanofibers have a large surface area, they are good to use as a drug or other biomolecular carrier to achieve efficient delivery. The tetracycline drug was delivered from polymer nanofibers using PLA, poly(ethylene–co–vinylacetate) (PEVA), and their composites [119]. Kenawy et al. evaluated the *in vitro* release kinetics of the drug from the electrospun polymer mats, and compared them with a commercially available drug delivery system. The polymer mats of PEVA and PLA/PEVA composites delivered the drug in a controlled fashion as compared to other materials used in this experiment, such as PLA and commercial drug delivery system (Actisite). With the overview of illustrated examples, nanoscience and nanotechnology will certainly provide a new tool for the delivery of biomolecules in a controlled fashion. Further, this will make it possible to do molecular-level therapeutic release to diseased parts.

5. CHALLENGES FOR NANOTECHNOLOGY

Among the challenges facing nanotechnologists for rapid progress to continue in the field are the necessary advances that must be made in several enabling technologies. There are mainly two aspects associated with identifying the critical issues of nanostructural synthesis and assembly in which the control of grain size and the composition of nanomaterials play prominent roles. Another one is the control of interfaces and distributions of the nanoparticles within the materials. The first aspect is inextricably linked; nevertheless, it is important to understand how to exercise separate control over the nucleation of the nanostructural building blocks and the growth of those components throughout the synthesis and assembly process. It is always appreciated to control and/or manipulate the size of matter at an ultrafine level, and to establish computational and mathematical approaches to direct the same in the right path. The latter issue highlights an ability to scale up synthesis and assembly strategies at low cost, large-scale production of nanostructure materials, and control the size and quality of interfaces. Thus, nanotechnology helps us to understand how the structure of biomaterials is connected to their properties at the nanolevel, and thereby enhances the design of biomaterials with special properties required for particular site applications. Sometimes, stability also needs to be considered significantly since nanostructures are often inherently unstable owing to their small constituent sizes and high chemical activity. Hence, a further challenge is to enhance the stability of these materials in terms of structural, thermal, chemical, and biological aspects. The relationship between the micro- and nanoscale properties of the biomaterials has been established to some extent. As a result, nanomaterials available today are mechanically stronger, have a larger surface area, and are more durable, reproducible, and surface modifiable to specific needs. Further investigations are needed on the biological performance of nanostructured biomaterials to understand their interactions with the body environment at molecular level, and to increase their potential as human health care substitutes.

6. CONCLUSIONS

Nature always presents challenges for materials scientists and engineers to design medical implants or scaffolds similar to natural tissue or organs with respect to their performance. With the aid of nanoscience and technology, considerable investigations on medical prosthetics have provided the opportunity to design materials at the nanolevel with all ideal characteristic features in order to function in a natural way in our body environment. Novel innovative technology for the fabrication of nanostructural biomaterials helps increase the possibility of realizing implants with enhanced quality, at both the *in vitro* and *in vivo* levels. Nanostructured biomaterials are a generous gift to humans, either as temporary or permanent substitutes for defective or degenerated tissues and organs. They represent potential breakthroughs in human health care, including the rejuvenation of bone graft materials, three-dimensional scaffolds for tissue engineering, and macromolecular delivery systems using nanoscale particles. Compared to microscale biomaterials, nanoscale biomaterials are relatively less available. Now, the trend is changing to nanoscale biomaterials in order to mimic the natural tissue functions after implantation. With the advancement of nanoscience and nanotechnology, it is expected to develop nanoscale biomaterials either in a monolithic or a composite form in human health care systems with all of the qualities of ideal biomaterials, which leads to an improved quality of human life.

GLOSSARY

Allograft Transplantation of tissue or organ between individuals of the same species, but of non-identical genetic.

Alloplastic graft The use of synthetic biomaterials to repair or replace a defective part of the body.

Artificial organ A medical device that replaces, in part or whole, the functions of the human organs.

Autograft Transplantation of tissue or organ from one part of a body to another part of the same body.

Bioactive material A material, which is capable to elicit an intended biological function.

Biocompatibility The natural ability of an implant material to perform with an appropriate host response in a specific application.

Biodegradability The susceptibility of an implant material to decompose by a living organism.

Bioimplant A scientific term representing the prosthesis made of biocompatible synthetic material.

Biomaterials Any synthetic materials which are compatible with living tissues upon implantation. They can be metals, ceramics, polymers, and their composites.

Bone A supportive rigid connective tissue consisting of collagen fibers impregnated with calcium based minerals in conjunction with metabolic cells.

Cell The fundamental structural and functional unit of all life, that is, composed of an outer membrane enclosing protoplasm and nucleus.

Collagen A fibrous connective tissue protein having triple helical structure that functions to hold tissues together.

Composite A material, which contains two or more different substances.

Compressive modulus The ratio of compressive stress to compressive strain below the proportional limit.

Compressive strength The maximum load sustained by a test specimen in a compressive test divided by the original area of the specimen.

Connective tissue Any of the tissues of the body that support the specialized elements or parenchyma.

Cortical bone The compact hard bone with functional osteons.

Drug Any substance other than food or water that is intended to be taken for the purpose of altering, sustaining or controlling the recipient's physical or mental state.

Enamel A hard white substance that covers the dentine of the crown of a tooth. It is the hardest substances in the body.

Encapsulation A process of applying a thin coating onto a substrate, which conforms to the external geometry of the article being encapsulated.

FDA An abbreviation for Food and Drug Administration, the agency formed in 1931 under the Department of Health Education and Welfare, USA. It deals with the safety and efficacy of products marketed for biomedical application.

Femur An orthopedic term of thigh bone, that is, the bone of upper leg.

Fracture The separation of a body part, characterized in either brittle or ductile.

Graft A transplant.

Growth factors A heterogeneous group of substances whose general purpose is to enhance the rate of healing.

Hardness The resistance of a plastic to indentation, compression and scratching.

Hard tissue The general term for calcified structures in the body, e.g., bone and tooth.

Hydroxyapatite An inorganic phase of calcified bone tissue. It is a type of calcium phosphate having a chemical composition of $Ca_{10}(PO_4)_6(OH)_2$.

Implant Any medical device made from single or multi phase materials intended for human health care applications.

Implantation A surgical procedure by which medical devices or prosthesis is placed, either temporarily or permanently, in human body.

Infection A disease process caused by the invasion and the damaging action of microorganisms like bacteria or virus.

In vitro A biological study performed outside the living body.

In vivo A biological study performed inside the living body.

Nanomaterials Materials composed of particles having nanometric tolerances; a nanometer is being equal to 10^{-9} m.

Organ A differentiated part of an organism adapted for a definite function.

Orthopedics The medical specialty concerned with the skeletal system.

Osteoblast Any one of the cells of mesenchymal origin concerned to the formation of bone tissue.

Osteoclast One of the large multinuclear cells found in association with the resorption of bone tissue.

Osteocyte A cell, which is responsible for the maintenance of bone tissue.

pH A numerical expression of the degree of acidity or alkalinity of an aqueous solution.

Polymer A large molecule formed by the combination of many similar smaller molecules i.e. monomers.

Porosity A ratio of the void volume to the total volume expressed in terms of percentage.

Prosthesis A device that replaces a limb, organ or tissue of the body.

Resorption Dissolution or removal of a substance.

Scaffold A biocompatible 3D structural constructs with or without cells intend as a temporary implantable tissue capable to be replaced by natural tissue.

Surface area The total area of exposed surface of a powder, form or textured material.

Tissue A collection of similar cells and the surrounding intercellular substances.

Tissue engineering The development and manipulation of laboratory grown molecules, cells, tissues, or organs to repair or replace the defective body parts.

Transplantation Surgical transfer of a tissue or organ from one pace to another.

Trauma Any injury caused by external force.

Tumor An uncontrolled growth of cells that serve no useful purpose, which may be benign or malignant.

Xenograft Transplantation of tissue or organ between two different species.

ACKNOWLEDGMENTS

Financial support of the NUS Research Project (R-398-000-002-112) is gratefully acknowledged.

REFERENCES

1. C. M. Agrawal and K. A. Athanasiou, *J. Biomed. Mater. Res. Appl. Biomater.* 38, 105 (1997).
2. C. M. Agrawal and R. B. Ray, *J. Biomed. Mater. Res.* 55, 141 (2001).
3. C. M. Agrawal, J. S. McKinney, D. Huang, and K. A. Athanasiou, in "Synthetic Bioabsorbable Polymers for Implants" (C. M. Agrawal, J. E. Parr, and S. T. Lin, Eds.) American Society for Testing and Materials, Philadelphia, PA, 2000.
4. E. S. Ahn, N. J. Gleason, A. Nakahira, and J. Y. Ying, *Nano Lett.* 1, 149 (2001).
5. H. Aoki, "Science of Medical Applications of Hydroxyapatite," p. 137. Takayama Press, Tokyo, 1991.
6. S. A. Athreya and D. C. Martin, *Sensors Actuators A—Phys.* 72, 203 (1999).
7. Y. H. Bae, K. M. Huh, Y. Kim, and K. J. Park, *J. Contr. Res.* 64, 3 (2000).
8. P. J. Barham and A. Keller, *J. Mater. Sci.* 20, 2281 (1985).
9. R. A. Barkhordar and J. R. Meyer, *Oral Surg.* 61, 201 (1986).
10. D. V. Bazile, C. Robert, P. Huve, T. Verrecchia, M. Marland, A. Frydman, M. Veillard, and T. Spenlehauer, *Biomater.* 13, 1093 (1992).
11. J. P. Berry, U.S. Patent 4,965,110, 1990.

12. M. J. Bissell and M. H. Barcellos-Hoff, *J. Cell. Sci. (Suppl.)* 8, 327 (1987).
13. J. Black and G. W. Hastings, "Handbook of Biomaterials Properties." Chapman and Hall, London, 1998.
14. J. Blaine, *Lancet* 2, 525 (1946).
15. A. S. Blawas and W. M. Reichert, *Biomateri.* 19, 595 (1998).
16. A. Bornat, U.S. Patent 4,689,186, 1987.
17. S. A. Bortz and E. J. Onesto, *Ceram. Bull.* 52, 898 (1973).
18. M. P. Bostrom and N. P. Camacho, *Clin. Orthop. Rel. Res.* S274 (1998).
19. P. Boutin, P. Christel, J. M. Dorlot, A. Meunier, A. Roquancourt, D. Blanquaert, S. Herman, I. Sedel, and J. Witvoet, *J. Biomed. Mater. Res.* 22, 1203 (1988).
20. R. Brotzman, in "WTEC Panel Report on Nanostructure Science and Technology: R&D Status and Trends in Nanoparticles, Nanostructured Materials, and Nanodevices" (R. W. Siegel, E. Hu, and M. C. Roco, Eds.), p. 122. WTEC, MD, 1999.
21. C. J. Buchko, L. C. Chen, Y. Shen, and D. C. Martin, *Polymer* 40, 7397 (1999).
22. C. J. Buchko, K. M. Kozloff, and D. C. Martin, *Biomateri.* 22, 1289 (2001).
23. C. J. Buchko, M. J. Slattery, K. M. Kozloff, and D. C. Martin, *J. Mater. Res.* 15, 231 (2000).
24. A. Bukar, P. Fassina, F. Greco, S. Piantelli, C. Piconi, and N. Zaghini, in "Bioceramics 3" (J. E. Hulbert and S. F. Hulbert, Eds.), p. 355. Rose-Hulman Institute of Technology Publication, Terre Haute, IN, 1990.
25. K. J. L. Burg, W. D. Holder, C. R. Culberson, R. J. Beiler, K. G. Greene, A. B. Loebsack, W. D. Roland, D. J. Mooney, and C. R. Halberstadt, *J. Biomater. Sci., Poly. Ed.* 10, 147 (1999).
26. M. M. Burgshoef and G. J. Vancso, *Adv. Mater.* 16, 1362 (1999).
27. I. Catelas, L. H. Yahia, O. Huk, and R. Marchand, in "Transactions of the 4th World Biomaterials Congress," Toronto, Canada, June 1996, p. 11.
28. S. Chand, *J. Mater. Sci.* 35, 1303 (2000).
29. K. D. Chesmel and J. Black, *J. Biomed. Mater. Res.* 29, 1089 (1995).
30. J. Chevalier, J. M. Drouin, and B. Cales, in "Low Temperature Ageing Behaviour of Zirconia Hip Joints Heads, Bioceramics" (L. Sedel and C. Ray, Eds.), p. 135. Elsevier, Amsterdam, 1997.
31. P. Christel, A. Meunier, and J. M. Dorlot, *Ann. NY Acad. Soc.* 523, 243 (1988).
32. P. Clark, P. Connolly, A. S. G. Curtis, J. A. T. Dow, and C. D. W. Wilkinsin, *Develop.* 108, 635 (1990).
33. I. C. Clarke and G. Willmann, in "Bone Implant Interface" (H. U. Cameron, Ed.). Mosby Year Book Publications, St.Louis, MO, 1994.
34. A. G. Coombes and J. D. Heckman, *Biomater.* 13, 217 (1992).
35. M. L. Cooper, J. F. Hansbrough, R. L. Spielvogel, R. Cohen, R. L. Bartel, and G. Naughton, *Biomater.* 12, 243 (1991).
36. S. Croteau, F. Rauch, A. Silvestri, and R. C. Handy, *Orthopaedics* 22, 686 (1999).
37. F. Z. Cui, C. Du, X. W. Su, X. D. Zhu, and N. M. Zhao, *Cells Mater.* 6, 31 (1996).
38. Y. Dashy and D. H. Reneker, *J. Electrostat*. 35, 115 (1995).
39. A. Daubree, *Compt. Rend. Acad. Sci. Paris* 32, 625 (1851).
40. J. R. Dees and J. E. Spruiell, *J. Appl. Poly. Sci.* 18, 1053 (1974).
41. K. DeGroot, *Tissue Eng.* 4, 337 (1998).
42. K. DeGroot, in "Chemistry of Calcium Phosphates" (T. Yammamuro and L. L. Hence, Eds.). CRC Press, Boca Raton, FL, 1990.
43. J. M. Deitzel, J. Kleinmyeyr, D. Harris, and N. C. Beck Tan, *Polymer* 42, 261 (2001).
44. R. P. Desjardins, *J. Prosthetic Dent.* 54, 374 (1985).
45. I. Dion, L. Bordenave, and F. Lefebvre, *J. Mater. Sci., Mater. Med.* 5, 18 (1994).
46. A. J. Domb, in "Polymeric Site Specific Pharmacotherapy," p. 243. Wiley, New York, 1994.
47. F. C. M. Driessens, M. H. Boltong, E. A. P. De Maeyer, R. Wenz, B. Nies, and J. A. Planell, *Biomater*. 23, 4011 (2002).
48. C. Du, F. Z. Cui, X. D. Zhu, and K. DeGroot, *J. Biomed. Mater. Res.* 44, 407 (1999).
49. V. A. Dubok, *Powder Metall. Met. Ceram.* N7-8, 69 (2000).
50. P. Ducheyene, in "Bioceramics: Material Characteristics Versus *In-vivo* Behavior" (P. Ducheyene and J. E. Lemons, Eds.), p. 523 Annals of the New York Academy of Science, New York, 1988.
51. P. Ducheyne, *J. Biomed. Mater. Res. Appl. Biomater.* A21, 219 (1987).
52. P. Ducheyne, in "Bone Bonding Biomaterials" (P. Ducheyne, T. Kokubo, and C. A. Van Blitter Swijk, Eds.). Reced Health Care Communications, 1992.
53. S. A. Edward, J. G. Nathaniel, N. Atsushi, and Y. Y. Jackie, in "MRS 1998 Fall Meetings" (D. Xiao and J. Ying, Eds.), p. 611. private communication, 1998.
54. H. W. Elgendy, M. E. Norman, A. R. Keaton, and C. T. Laurencin, *Biomater*. 14, 263 (1993).
55. K. L. Elias, R. L. Price, and T. S. Webster, *Biomater*. 23, 3279 (2002).
56. L. G. Ellies, J. M. Carter, J. R. Natiella, J. D. B. Featherstone, and D. G. A Nelson, *J. Biomed. Mater. Res.* 22, 137 (1985).
57. J. C. Elliot, in "Structure and Chemistry of the Apatites and other Calcium Orthophosphates," p. 111. Elsevier, Amsterdam, 1994.
58. C. Emile, D. Bazile, F. Herman, M. Helene, and M. Veillard, *Drug Del*. 3, 187 (1996).
59. I. Engelberg and J. Kohn, *Biomater.* 12, 292 (1991).
60. W. Eysel and D. M. Roy, *J. Cryst. Growth* 20, 245 (1973).
61. L. Feng, S. Li, H. Li, J. Zhai, Y. Song, L. Jiang, and D. Zhu, *Angew Chem., Int. Ed.* 41, 1221 (2002).
62. R. G. Flemming, C. J. Murphy, G. A. Abrams, S. L. Goodman, and P. F. Nealey, *Biomater.* 20, 573 (1999).
63. H. Fong and D. H. Reneker, in "Structure Formation in Polymeric Fibers" (D. R. Salem, Ed.), p. 225. Hanser, Munich, 2001.
64. A. Formhals, U.S. Patent 1,975,504, 1934.
65. A. Formhals, U.S. Patent 2,077,373, 1937.
66. A. Formhals, U.S. Patent 2,159,416, 1939.
67. A. Formhals, U.S. Patent 2,160,962, 1939.
68. A. Formhals, U.S. Patent 2,187,306, 1940.
69. A. Formhals, U.S. Patent 2,323,025, 1940.
70. A. Formhals, U.S. Patent 2,349,950, 1944.
71. R. C. Garvie, C. Urbani, D. R. Kennedy, and J. C. Mc Neuer, *J. Mater. Sci.* 19, 3224 (1984).
72. T. N. Gerhart, A. A. Renshaw, R. L. Miller, R. J. Noecker, and W. C. Hayes, *J. Biomed. Mater. Res.* 23, 1 (1989).
73. P. W. Gibson, H. L. S. Gibson, and D. Rivin, *Colloids Surf. A: Physicochem. Eng. Aspects* 187–188, 469 (2001).
74. P. W. Gibson, H. L. S. Gibson, and D. Rivin, *AIChE J.* 45, 190 (1999).
75. P. W. Gibson and H. L. S. Gibson, U.S. Army Soldier and Biological Chemical Command Technical Report, Natick/TR99/016L, 1996.
76. R. A. Giordano, B. M. Wu, S. W. Borland, L. G. Cima, E. M. Sachs, and M. J. Cima, *J. Biomater. Sci., Poly. Ed.* 8, 63 (1996).
77. S. L. Goodman, P. A. Sims, and R. M. Albrecht, *Biomater.* 17, 2087 (1996).
78. T. Gorner, R. Gref, D. Michenot, F. Sommer, M. Tran, and E. Dellacherie, *J. Contr. Res.* 57, 259 (1999).
79. A. R. Greenberg, *J. Biomed. Mater. Res.* 12, 922 (1978).
80. R. Gref, Y. Minamitake, M. T. Peracchia, V. Trubetskoy, V. Torchilin, and R. Langer, *Science* 263, 1600 (1994).
81. R. Gref, in "Synthesis, Functionalization and Surface Treatment of Nanoparticles" (M. I. Baraton, Ed.). American Scientific, CA, 2003.
82. L. G. Griffith, *Acta Mater.* 48, 263 (2000).
83. U. Gross, H. J. Schmitz, and V. Struz, in "Bioceramics: Material Characteristics Versus *In-vitro* Behavior" (P. Ducheyne and

J. E. Lemons, Eds.), Vol. 523, p. 211. Annals of the New York Academy of Science, 1988.
84. P. Guiot and P. Couvreur, "Polymeric Nanoparticles and Microspheres." CRC Press, Boca Raton, FL, 1986.
85. Y. Han, K. Xu, G. Montay, T. Fu, and J. Lu, *J. Biomed. Mater. Res.* 60, 511 (2002).
86. P. Hanarp, D. Sutherland, J. Gold, and B. Kasemo, *Nanostruct. Mater.* 12, 429 (1999).
87. J. S. Hanker and B. L. Giammara, "MRS Symposium Proceedings of Biomedical Materials and Devices" (T. Kanazawa, Ed.), Materials Research Society, Pittsburgh, PA, 1989, p. 110.
88. M. F. Harmand, *Clin. Mater.* 11, 145 (1992).
89. L. D. Harris, B. S. Kim, and D. J. Mooney, *J. Biomed. Mater. Res.* 42, 396 (1998).
90. T. Hattori, Y. Iwadate, and T. Kato. *Adv. Ceram. Mater.* 3, 426 (1988).
91. H. Hayashi, A. Uchida, H. Hamada, H. Yoshikawa, Y. Shinto, and K. Ono, *Arch. Orthop. Trauma Surg.* 112, 1 (1992).
92. J. D. Helmer and T. D. Driskell, "Research on Bioceramics, Symposium on Use of Ceramics as Surgical Implants," Clemson University, SC, 1969.
93. L. L. Hench, *J. Am. Ceram. Soc.* 81, 1705 (1998).
94. L. L. Hench, *J. Am. Ceram. Soc.* 74, 1484 (1991).
95. T. V. How, U.S. Patent 4,552,707, 1985.
96. L. Huang, R. P. Apkarian, and E. L. Chaikof, *Scanning* 23, 372 (2001).
97. S. R. Hulbert, J. C. Bokros, L. L. Hench, J. Wilson, and J. Heimke, in "High Tech Ceramics" (Vincenzini, Ed.), p. 189. Elsevier, Amsterdam, 1987.
98. J. C. Hulteen, X. Chen, C. K. Chambliss, and C. R. Martin, *Nanostruct. Mater.* 9, 133 (1997).
99. C. D. Hummer, R. H. Rothman, and W. J. Hozack, *J. Arthr.* 10, 848 (1995).
100. J. A. Hunt, R. L. Williams, S. M. Tavakoli, and S. T. Riches, *J. Mater. Sci., Mater. Med.* 6, 813 (1995).
101. Y. Y. Huse, J. D. Gresser, D. J. Trantolo, C. M. Lyons, P. R. Gangadharan, amd D. L. Wise, *J. Biomed. Mater. Res.* 35, 107 (1997).
102. D. W. Hutmacher, A. Kirsch, K. L. Ackermann, M. B. Huerzeler, and M. Dent, *Int. J. Periodont. Restorative Dent.* 21, 49 (2001).
103. F. Ignatious and J. M. Baldoni, PCT/US01/02399, 2001.
104. L. Illum, S. S. Davis, S. Muller, E. Mak, and P. West, *Life Sci.* 40, 367 (1987).
105. F. D. Ingrahm, E. Alexander, and D. D. Matson, *New Eng. J. Med.* 236, 362 (1947).
106. International Standard Organization, ISO TC150/SC, DIS 13356, 1995.
107. K. Ioku, S. Yamauchi, H. Fujimori, S. Goto, and M. Yoshimura, *Solid State Ionics* 151, 147 (2002).
108. A. Ito, T. Tateishi, S. Niwa, and S. Tange, *Clin. Mater.* 12, 203 (1993).
109. E. B. Jaffe, Abstracts of the Literature on Synthesis of Apatites and Some Related Phosphates, *U.S. Geol. Surv., Circular* 135, 947 (1974).
110. M. Jarcho, C. H. Bolen, M. B. Thomas, J. Bobick, J. E. Kay, and R. H. Doremus, *J. Mater. Sci.* 11, 2027 (1976).
111. M. Jarcho, *Clin. Orthop. Rel. Res.* 151, 259 (1981).
112. H. J. Jin, S. Fridrikh, G. C. Rutledge, and D. Keplan, *Abst. Papers, Amer. Chem. Soc.* 224, 408 (2002).
113. P. D. Johnson, *J. Electrochem. Soc.* 108, 159 (1961).
114. H. Kaczmarek, L. A. Linden, and J. F. Rabek, *J. Appl. Poly. Sci.* 60, 2341 (1996).
115. T. Kanazawa, "Inorganic Phosphate Materials." Elsevier, Amsterdam, 1989.
116. F. S. Kaplan, W. C. Hayes, T. M. Keaven, A. Boskey, T. A. Einhorn, and J. P. Iannotti, in "Orthopedic Basic Science" (S. R. Simon, Ed.), p. 127. American Academy of Orthopedic Surgeons, Columbus, OH, 1994.
117. B. Kasemo and J. Lausmaa, *CRC Crit. Rev. Biocomp.* 2, 335 (1986)
118. M. I. Kay, R. A. Young, and A. S. Posner, *Nature* 204, 1051 (1964).
119. E. R. Kenawy, G. L. Bowlin, K. Mansfield, J. Layman, D. G. Simpson, E. H. Sanders, and G. E. Wnek, *J. Contr. Rel.* 81, 57 (2002).
120. G. Kickelbick and U. Schubert, in "Synthesis, Functionalization and Surface Treatment of Nanoparticles" (M. I. Baraton, Ed.). American Scientific Publishers, Stevenson Ranch, CA, 2003.
121. T. Kikuki, M. Ohgaki, M. Katsura, S. Nakamura, K. Hashimoto, Y. Toda, S. Udagawa, and K. Yamashita, *Biomater.* 24, 941 (2003).
122. G. C. Koumoulidis, T. C. Vaimakis, and T. Sdoukos, *J. Am. Ceram. Soc.* 84, 1203 (2001).
123. H. H. Kung and E. I. Ko, *Chem. Eng. J.* 64, 203 (1994).
124. C. T. Laurencin, A. M. A. Ambrosia, M. D. Borden, and J. A. Cooper, Jr., *Ann. Rev. Biomed. Eng.* 1, 19 (1999).
125. R. Z. Legeros, *Adv. Dent. Res.* 2, 164 (1988).
126. R. Z. LeGeros, in "Biological and Synthetic Apatites" (P. W. Brown and B. Constantz, Eds.), p. 3. CRC Press, Boca Raton, FL, 1994.
127. S. R. Levitt, P. H. Crayton, E. A. Monroe, and R. A. Condrate, *J. Biomed. Mater. Res.* 3, 683 (1969).
128. J. Li, Y. Liu, L. Hermansson, and R. Soremark, *Clin. Mater.* 12, 197 (1993).
129. W. J. Li, C. T. Laurencian, E. J. Caterson, R. S. Tuan, and F. K. Ko, *J. Biomed. Mater. Res.* 60, 613 (2002).
130. G. J. Liu, J. F. Ding, L. J. Qiao, A. Guo, B. P. Dymov, J. T. Gleeson, T. Hashimoto, and K. Saijo, *Chem—A Eur. J.* 5, 2740 (1999).
131. Q. Liu, J. R. de Wijn, and C. A. Van Blitterswijk, *Biomater.* 18, 1263 (1997).
132. Q. Liu, J. R. de Wijn, and C. A. Van Blitterswijkrek, *J. Biomed. Mater. Res.* 40, 257 (1998).
133. Y. Liu, W. Wang, Y. Zhan, C. Zhang, and G. Wang, *Mater. Lett.* 56, 496 (2002).
134. Lord Rayleigh, *London, Edinburgh and Dublin Phil. Mag. J.* 44, 184 (1882).
135. L. Lu and A. G. Mikos, *MRS Bull.* 11, 28 (1996).
136. P. X. Ma and R. Zhang, *J. Biomed. Mater. Res.* 46, 60 (1999).
137. M. Manso, S. Ogueta, P. H. Fernandez, L. Vazquezl, M. Langlet, and J. P. G. Ruiz, *Biomateri.* 23, 3985 (2002).
138. C. R. Martin, *Science* 266, 1961 (1994).
139. C. R. Martin, *Chem. Mater.* 8, 1739 (1996).
140. G. E. Martin, I. D. Cockshott, and F. J. T. Fildes, U.S. Patent 4,044,404, 1977.
141. G. E. Martin, I. D. Cockshott, and F. J. T. Fildes, U.S. Patent 4,878,908, 1989.
142. E. Mathiowitz, C. Amoto, and R. Langer, *Polymer* 31, 547 (1990).
143. J. Matsumoto, Y. Nakada, K. Sakurai, T. Nakamura, and Y. Takahashi, *Int. J. Pharm.* 185, 93 (1999).
144. J. A. Matthews, G. E. Wnek, D. G. Simpsm, and G. L. Bowlin, *Biomacromol.* 3, 232 (2002).
145. J. W. McCauley and R. Roy, *Am. Miner.* 59, 947 (1974).
146. A. G. Mikos, Y. Bao, L. G. Cima, D. E. Ingber, J. P. Vacanti, and R. Langer, *J. Biomed. Mater. Res.* 27, 183 (1993).
147. A. G. Mikos, G. Saraikihos, S. M. Leite, J. P. Vacanti, and R. Langer, *Biomater.* 14, 323 (1993).
148. S. Miyamoto and K. Takaoka, *Ann. Chir. Gynaecol.* 82, 69 (1993).
149. S. Moghimism, C. Porter, S. S. Davis, and L. Illum, *Biochem. Biophys. Res. Commun.* 177, 861 (1991).
150. H. Monma and T. Kamiya, *J. Mater. Sci.* 22, 4247 (1987).
151. H. Monma, T. Takahashi, H. Ushio, S. Soeda, and T. Kiyosawa, *Gypsum & Lime* 22, 25 (1988).
152. H. Monma, S. Ueno, and U. Kamiya, *J. Chem. Tech. Biotechnol.* 31, 15 (1981).
153. D. J. Mooney, D. F. Baldwin, N. P. Suh, J. P. Vacanti, and R. Langer, *Biomater.* 17, 1417 (1996).
154. A. Mortier, J. Lemaiter, L. Rodrique, and P. G. Rouxhet, *J. Am. Ceram. Soc.* 78, 215 (1989).

155. C. B. Murray, D. J. Noris, and M. G. Bawendi, *J. Am. Chem. Soc.* 115, 8706 (1993).
156. R. Murugan and K. P. Rao, *J. Biomater. Sci., Poly. Ed.* (2003).
157. R. Murugan and K. P. Rao, *Macromol. Res.* 11, 14 (2003).
158. R. Murugan and K. P. Rao, *Polym. Bull.* 49, 402 (2003).
159. R. Murugan and K. P. Rao, *Trends Biomater. Art. Org.* 16, 43 (2002).
160. N. Nagai and N. Takeshita, *J. Dent. Met.* 25, 581 (1982).
161. S. Nakamura, T. Tsobe, and M. Senna, *J. Nanopart. Res.* 3, 57 (2001).
162. L. S. Nichter, M. Yazdi, Y. Kosari, R. Sridjaja, E. Ebramzadeh, and N. E. Nimni, *J. Craniofac. Surg.* 3, 63 (1992).
163. R. R. Nielson, T. Karring, and S. Gogolewski, *Acta Orthop. Scand.* 63, 66 (1992).
164. I. D. Norris, M. M. Shaker, F. K. Ko, and A. G. Mac Diarmid, *Synth. Metals* 114, 109 (2000).
165. P. T. Ohara and R. C. Buck, *Exp. Cell. Res.* 121, 235 (1979).
166. T. Ondarcuhu and C. Joachim, *Europhys. Lett.* 42, 215 (1998).
167. J. F. Osborn and H. Newesely, *Biomater.* 1, 108 (1980).
168. Y. Ota and T. Iwashita, *J. Am. Ceram. Soc.* 81, 1665 (1998).
169. Z. Panagi, P. Beletsi, G. Evangelatos, E. Livaniou, D. S. Ithakissios, and K. Avgoustakis, *Int. J. Pharm.* 221, 143 (2001).
170. R. N. Panda, M. F. Hsieh, R. J. Chung, and T. S. Chin, *J. Phys. Chem. Solids* 64, 193 (2003).
171. A. Park and L. G. Griffith-Cima, *J. Biomed. Mater. Res.* 31, 117 (1996).
172. J. B. Park, "Biomaterials Science & Engineering." Plenum, New York, 1984.
173. W. Paul and C. P. Sharma, *Trends Biomater. Art. Org.* 14, 37 (2001).
174. A. Peigney, C. H. Laurent, and H. Rousset, *Key Eng. Mater.* 132–136, 743 (1997).
175. M. T. Peracchia, R. Gref, Y. Minamitake, A. Domb, N. Lotan, and R. Langer, *J. Contr. Res.* 57, 259 (1999).
176. A. S. Posner, *Clin. Orthop. Rel. Res.* 200, 87 (1985).
177. POST, Preliamientary Office of Science and Technology, Vaccines and Their Future Role in Public Health, London, 1995, POST Report 66.
178. R. L. Price, M. C. Waid, K. M. Haberstroh, and T. J. Webster, *Biomater.* 24, 1877 (2003).
179. D. A. Puleo, L. A. Holleran, R. H. Doremus, and R. Bizios, *J. Biomed. Mater. Res.* 25, 711 (1991).
180. S. Ramakrishna, J. Mayer, E. Wintermantel, and K. W. Leong, *Composites Sci. Tech.* 61, 1189 (2001).
181. A. L. Ray, M. Vert, J. Gauthier, and G. Benoit, *Int. J. Pharm.* 106, 201 (1994).
182. H. Redhead, S. S. Davis, and L. Illum, *J. Controlled Rel.* 70, 353 (2001).
183. D. H. Reneker and I. Chun, *Nanotechnol.* 7, 216 (1996).
184. M. Z. Rong, L. J. Qiu, M. Q. Zhang, and K. Friedrich, *Eur. Poly. J.* 38, 1573 (2002).
185. M. Z. Rong, M. Q. Zhang, X. Y. Zheng, H. M. Zeng, R. Walter, and K. Friedrich, *Polymer* 42, 167 (2001).
186. T. Sakou, *Bone* 22, 591 (1998).
187. J. Saphiro, S. Reisner, R. Lichtenberg, and R. Meltzer, *J. Am. Coll. Cardiol.* 16, 1603 (1990).
188. J. M. Schnur, *J. Controlled Rel.* 28, 3 (1994).
189. A. G. Scopelianos, U.S. Patent 5,522,879, 1996.
190. Y. Shikinami and M. Okuno, *Biomater.* 20, 859 (1998).
191. Y. M. Shin, M. M. Hohman, M. P. Brenner, and G. C. Rutledge, *Polymer* 42, 9955 (2001).
192. R. W. Siegel, *Sci. Amer.* 275, 42 (1996).
193. R. W. Siegel, U.S. Patent 5,128,081, 1994.
194. R. W. Siegel, U.S. Patent, 5,320,800, 1994.
195. A. R. Singhal, C. M. Agrawal, and K. A. Athanasiou, *Tissue Eng.* 2, 197 (1996).
196. D. K. Smith, in "Calcium Phosphate Apatites in Nature" (P. W. Brown and B. Constantz, Eds.), p. 23. CRC Press, Boca Raton, FL, 1994.
197. E. Solheim, E. M. Pinholt, G. Bang, and E. Sudmann, *J. Neurosurg.* 76, 275 (1992).
198. M. D. Stenoien, W. J. Drasler, R. J. Scott, and M. L. Jenson, U.S. Patent 5,866,217, 1999.
199. W. Tan, R. Krishnaraj, and T. A. Desai, *Tissue Eng.* 7, 203 (2001).
200. G. Taylor, *Proc. R. Soc.* A280, 383 (1964).
201. H. Terai, D. Hannouche, E. Ochoa, Y. Yamano, and J. P. Vacanti, *Mater. Sci. Eng. C* 20, 3 (2002).
202. R. C. Thomson, M. C. Wake, M. J. Yaszemski, and A. G. Mikos, *J. Biomater. Sci., Polymer Ed.* 7, 83 (1995).
203. R. C. Thomson, M. J. Yaszemski, and M. G. Mikos, in "Principles of Tissue Engineering" (R. P. Lanza, R. Langer, and W. L. Chick, Eds.), p. 263. R. S. Landes Co., Austin, TX, 1997.
204. R. C. Thomson, M. J. Yaszemski, and J. M. Powers, in "Biomaterials for Drugs and Cell Delivery" (A. G. Mikos, Ed.), p. 33. Materials Research Society, Pittsburgh, 1994.
205. R. C. Thomson, M. J. Yaszemski, and J. M. Powers, *J. Biomater. Sci., Polymer Ed.* 7, 23 (1995).
206. A. Toni, C. G. Lewis, and A. Sudanese, *J. Arthroplasty* 9, 435 (1994).
207. M. Toriyama, A. Ravaglioli, C. Krajewski, E. Galasse, G. Roncari, and A. Piancastelli, *J. Mater. Sci.* 30, 3216 (1995).
208. E. Tsuruga, H. Takita, H. Itoh, Y. Wakisaka, and Y. Kuboki, *J. Biochem.* 121, 317 (1997).
209. A. Uchida, S. M. L. Nade, E. L. McCarthey, and W. Ching, *J. Bone Joint Surg.* 66B, 269 (1984).
210. A. Van Sliedredgt, A. M. Radder, K. de Groot, and C. A. Van Blitterswijk, *J. Mater. Sci., Mater. Med.* 3, 365 (1992).
211. C. C. P. M. Verheyen, J. R. de Wijn, C. A. Van Blitterswijk, K. de Groot, and P. M. Rozing, *J. Biomed. Mater. Res.* 27, 433 (1993).
212. S. Vijaya and H. Verma, *Mater. Lett.* 56, 827 (2002).
213. R. Walter, K. Friedrich, V. Privalko, and A. J. Savadori, *J. Adhesion* 64, 87 (1997).
214. K. Wang, C. H. Thomas, K. E. Healy, and G. Nuber, *Polymer* 36, 837 (1995).
215. R. Z. Wang, F. Z. Cui, H. B. Lu, H. B. Wen, C. L. Ma, and H. D. Li, *J. Mater. Sci. Lett.* 14, 490 (1995).
216. Y. Wang, S. Serrano, and J. J. Santiago-Aviles, *Synth. Metals* 9567, 1 (2002).
217. T. J. Webster, R. W. Siegel, and R. Bizios, *Biomater.* 20, 1221 (1999).
218. T. J. Webster, R. W. Siegel, and R. Bizios, *Biomater.* 21, 1803 (2000).
219. T. J. Webster, R. W. Siegel, and R. Bizios, *Nanostuct. Mater.* 12, 983 (1999).
220. Werner, *Gerhard' Grundr* 281 (1786).
221. K. Whang, D. C. Tsai , E. K. Nam, M. Aitken, S. M. Sprague, and P. K. Patel, *J. Biomed. Mater. Res.* 42, 491 (1998).
222. G. M. Whitesides and B. Grzybowski, *Science* 295, 2418 (2002).
223. M. S. Widmer, P. K. Gupta, L. Lu, R. K. Meszlenyi, G. R. D. Evans, K. Brandt, T. Savel, A. Gurlek, C. W. Patrick, and A. G. Mikos, *Biomater.* 19, 1945 (1998).
224. D. F. Williams and J. Cunningham, "Materials in Clinical Dentistry." Oxford University Press, Oxford, U.K. 1979.
225. D. F. Williams, "Definition in Biomaterials." Elsevier, Amsterdam, 1987.
226. P. Wiltzius, WTEC Panel Report on Nanostructure Science and Technology: R&D Status and Trends in Nanoparticles, Nanostructured Materials, and Nanodevices (R. W. Siegel, E. Hu, and M. C. Roco, Eds.), WTEC, MD, 1999, p. 119.
227. T. Yammamuro, L. L. Hench, and J. Wilson, in "Handbook of Bioactive Ceramics," Vol. 2. CRC Press, Boca Raton, FL, 1990.
228. F. Yang, R. Murugan, S. Ramakrishna, X. Wang, Y. Ma, and S. Wang, *Biomater.* (submitted, 2003).

229. Y. Yang and J. L. Ong, *J. Mater. Sci. Lett.* 21, 67 (2002).
230. Y. Yokayama, M. Toriyama, Y. Kawamto, T. Suzuki, K. Nishizawa, R. Nagata, and R. Mucalo, *Chem. Lett.* 25, 91 (1996).
231. H. Yoshimoto, Y. M. Shin, H. Terai, and J. P. Vacanti, *Biomater.* 24, 2077 (2003).
232. Z. Zang, M. F. Dunn, T. D. Xiao, A. P. Tomsia, and E. Saiz, *Mater. Res. Soc. Symp. Proc.* 7.5, 703 (2002).
233. J. Zeleng, *Phys. Rev.* 10, 1 (1917).
234. S. Zhang and K. E. Consalves, *J. Mater. Sci., Mater. Med.* 8, 25 (1997).
235. H. Y. Zhu and G. Q. Lu, in "Nanophase and Nanocomposite Materials IV" (S. Komarnerni, Ed.), p. 703. Materials Research Society, PA, 2001.
236. H. Y. Zhu and G. Q. Lu, Patent WO 01/74713 AI, Oct. 11, 2001.

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanostructured Bipolar Organic Polymers

Antonio Cravino, Helmut Neugebauer, N. Serdar Sariciftci

Johannes Kepler University Linz, Linz, Austria

CONTENTS

1. INTRODUCTION

During the last decades, the research activity of many chemists, physicists, and material scientists has been increasingly committed to the synthesis, understanding, and applicational use of molecules with extended π-electron systems. This work not only has allowed the development of new exciting concepts, but has also boosted the utilization of organic materials like conjugated polymers [1–4] and fullerene derivatives [5–12] in fields of application that have been so far reserved to inorganic semiconductors and metals. For example, after a period of academic research, polymeric films for antistatic coatings as well as polymeric light-emitting diodes and displays have entered the market [13–16]. Moreover, the discovery of a photoinduced electron transfer from nondegenerate ground-state conjugated polymers onto fullerene derivatives [17, 18] fueled up the research on organic polymeric photovoltaics (PV). Indeed, conjugated polymers and fullerenes became potentially suitable for the cost-effective fabrication of flexible, lightweight, and large-area photovoltaic devices and photodetectors [19, 20].

Organic solar cells of the first generation were p/n-heterojunction diodes. These devices were heterojunctions between organic thin films, one consisting of an electron-donating molecule (p-type layer), and the second of an electron-accepting molecule (n-type layer), sandwiched between two metal contacts with different workfunctions. These flat heterojunctions showed rectification behavior as well as photovoltaic effects [21–27]. However, this class of devices suffered limited interaction between the p and n components, which is in them confined to a flat geometrical interface. A breakthrough has been achieved by the direct blending of a soluble p-type conjugated polymer and C_{60} fullerene, or other fullerene derivatives, as an n-type component [27, 28]. In these devices, called "bulk-heterojunction" solar cells, the effective interaction between the donor and the acceptor components can take place in the entire bulk of the photoactive layer. As such, a more efficient charge generation was attained, which in turn dramatically enhanced the short circuit current J_{SC} values, resulting in much more efficient photovoltaic devices. However, the performance of bulk-heterojunction solar cells is still affected by some critical factors. Among them are the following.

- The solid-state miscibility between the conjugated polymer and the fullerene components plays a crucial role in determining the final morphology of the blend. Phase separation and clustering of the fullerene usually occur, reducing the effective donor/acceptor interfacial area as compared to the ideal maximum value [29].
- Clustering effects, leading to a "void" between p- as well as n-type domains, are dramatically detrimental for the transport of charges [29]. In the bulk-heterojunction architecture, a balanced transport of both photogenerated electrons and holes would be highly important. For an efficient transport of holes through the donor polymer matrix and of electrons through the acceptor fullerene phase, a truly bicontinuous and nanoscaled interpenetrating network is required.

As such, it is indeed appealing to think of a single material with intrinsic donor–acceptor properties, as well as being capable of both electron and hole transport (p/n-type material).

An elegant approach to such a target is offered by the design and synthesis of p-type conjugated backbones (donor *cable*) bearing directly grafted or tethered acceptor groups such as fullerene moieties (acceptor *cable*). In this way, the

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (615–628)

secondary structure of the macromolecule possibly might be dictated at a nanoscale level by its primary structure (double-cable polymer). The effective donor/acceptor interfacial area will be maximized, and phase separation and clustering phenomena should be prevented too. In addition, the interaction between the donor conjugated backbone and the acceptor moieties might be finely tuned by selecting the chemical structure (nature and length) of their connecting fragment [30]. Basically, the realization of effective double-cable polymers will bring the p/n-heterojunction concept at a nanoscale, molecular level. In this chapter, after a brief introduction to the photophysics of conjugated polymer:fullerene composites and to the bulk-heterojunction photovoltaic devices constructed using them, the developments of the design, characterization, and preliminary application of nanostructures consisting of a conjugated backbone bearing fullerene moieties are reviewed. Even though the search for electron-accepting components alternative to fullerenes has been pursued [31–36], the results so far obtained indicate that fullerenes are specially suitable for organic photovoltaics [27, 29, 37, 38].

2. CONJUGATED POLYMERS AND FULLERENES: PHOTOINDUCED ELECTRON TRANSFER

Most nondegenerate ground-state conjugated polymers in their photoexcited states, in which electrons are promoted to the antibonding π^* band, are electron donors. The stability of the so-called quasiparticles on conjugated polymers (e.g., radical cations such as positive polarons) [39, 40] suggested their use in combination with electron acceptor molecules. In 1992, the spectroscopic and photophysical properties of blends composed of poly((2–methoxy–5–(2′–ethylhexoxy)–*p*–phenylene) vinylene) or poly(3–octylthiophene) (MEH–PPV and P3OT, respectively; see Chart 1) and C_{60} [17, 18] were independently reported by Sariciftci et al. and Morita et al. While no ground-state interaction between the two components was observed, a strong quenching of the MEH–PPV luminescence indicated the existence of a subpicosecond photoinduced electron transfer [41]. The same behavior was observed later in a number of conjugate polymer/fullerene composites. Very recently, the transfer of electrons from photoexcited conjugated polymers onto 1–(3–methoxycarbonyl)–propyl–1–1–phenyl–(6, 6)C_{61} (PCBM, see Chart 1) has been studied by pump–probe measurements with the unprecedented time resolution of 10 fs [42]. The photoinduced electron transfer time was resolved with about 40 fs after an initial Kasha relaxation [42]. This high forward transfer rate results in a very efficient process, with a quantum yield that approaches unity. Finally, direct and definitive evidence for an electron transfer from photoexcited conjugated polymers onto fullerenes, leading to long-living charged states, was obtained by steady-state light-induced electron spin resonance (LESR) experiments. As an example, upon illumination of MDMO–PPV/PCBM composites (MDMO–PPV stands for poly((2–methoxy–5–(3, 7–dimethyloctyloxy)–*p*–phenylene) vinylene); see Chart 1), two light-induced electron spin resonance signals can be resolved: one at $g = 2.0026$, assigned to positive polarons on the conjugated backbone, and one at $g = 1.9997$, assigned to $PCBM^{\bullet -}$ radical anions [43, 44].

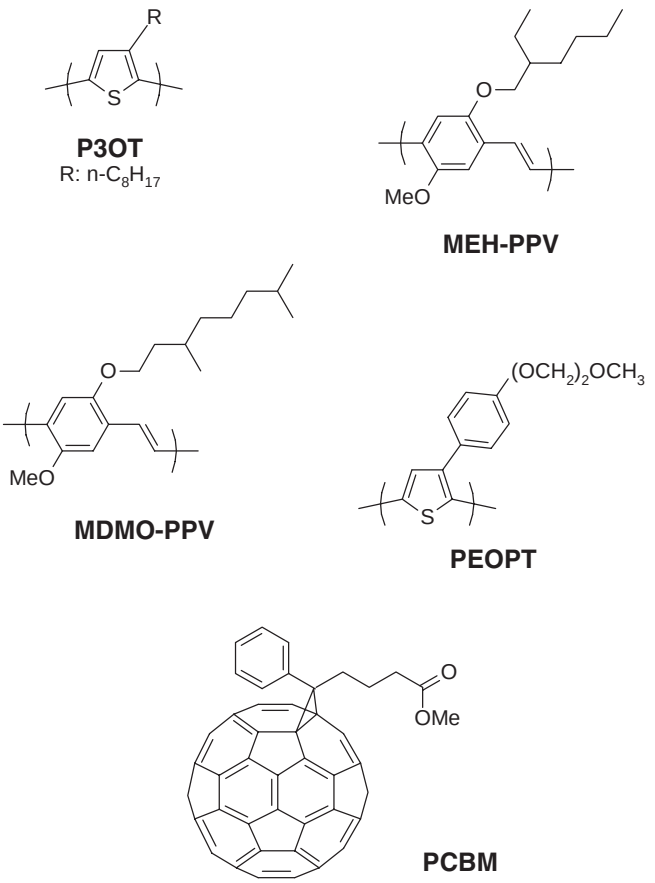


Chart 1

The long-lived nature of the charge carriers photogenerated within conjugated polymer:fullerene composites allows their collection at electrodes via diffusion (as well as field-induced migration), and their use to provide electrical power to an external utilizer. As such, these materials are rather attractive for PV solar energy conversion, and have been used for the preparation of plastic solar cells, as described in the next section.

3. ORGANIC BULK HETEROJUNCTIONS

In this section, we introduce the bulk-heterojunction solar cells, with the only aim to clarify the relationship between the morphology of the photoactive layer and the device's performance. Reference [27] gives a general and updated overview on organic, plastic PV. A discussion of the working principle and origin of the open-circuit voltage (V_{OC}) of conjugated polymer/fullerene bulk-heterojunction solar cells can be found in [37] and [45].

Extensive literature exists on the realization of bilayer photovoltaic elements based on small organic molecules, conjugated polymers, and fullerenes [21–26, 46, 47]. Although, as mentioned above, the internal quantum efficiency of photoinduced charge separation approaches unity for each

donor/acceptor pair, the power conversion efficiency of these devices is limited by the fact that only a small fraction of impinging photons is absorbed near the heterojunction interface. Charge separation occurs only at the flat geometrical interface, within the exciton diffusion length. Photoexcitations created far away from the heterojunction undergo recombination prior to reaching the electrodes. To overcome these problems, bulk-heterojunction solar cells, whose photoactive layer consists of a conjugated polymer/fullerene blend, have been designed [27, 28].

3.1. Morphology of Bulk Heterojunctions

Recently, it has been demonstrated that the efficiency of bulk-heterojunction plastic solar cells can be improved by manipulating the *morphology* within the photoactive blend [29]. Conjugated polymer:fullerene composites are, in fact, juxtapositions of domains with different compositions. Figure 1 shows images of the surfaces of MDMO–PPV/PCBM blends spun cast from a toluene solution (a) and a chlorobenzene solution (b), respectively, obtained by atomic force microscopy (AFM). The images show that a much more homogeneous mixing of the two components is obtained if chlorobenzene is used as the processing solvent. Recent morphological studies, performed by comparing AFM surface images and cross-sectional transmission electron microscopy (TEM) images of these blends, clearly corroborate this observation [48]. The impact of these two highly different morphologies to the PV device performance was studied by fabricating solar cells under identical conditions, except for the choice of the solvent. The open-circuit photovoltages for the two types of devices are nearly identical, while the device fabricated from the chlorobenzene solution exhibits more than a twofold enhancement in the J_{SC} as compared to the device cast from the toluene solution (5.25 versus 2.33 $mA \cdot cm^{-2}$). This high J_{SC} value, combined with a fill factor (FF) as high as 0.61, results in a power efficiency η of 2.5% under AM1.5 illumination [49]. The optical absorption spectra of the active layers are almost identical in the two cases, while the incident photons to current efficiency (IPCE, also called external quantum efficiency, EQE) of the device made from chlorobenzene is higher over the whole absorption spectrum. It is therefore clear that the difference in the device efficiencies must be related to the different morphologies shown in Figure 1. Shrinking both the interpenetrating two phases' domain size leads to an increase of the donor–acceptor interfacial area, as well as a decrease of the spatial separation between fullerene domains. Indeed, the increase in efficiency also relies on the increased mobility of the charge carriers of both signs, as indicated by field-effect measurements on the single components, as well as on the composite, spun cast from different solvents [50].

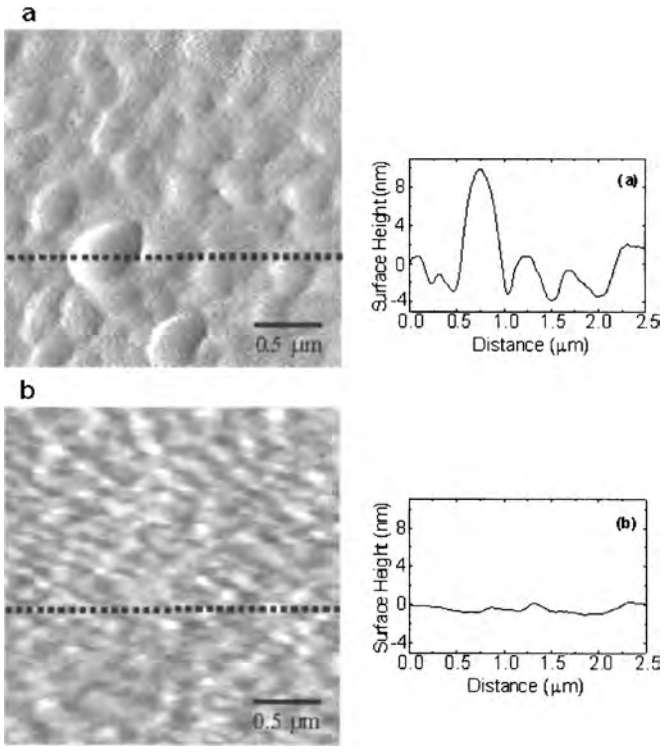


Figure 1. AFM images showing the surface morphology of MDMO–PPV/PCBM (1:4-w) blend films with thickness of about 100 nm and the corresponding cross section. (a) Film spun cast from solution in toluene. (b) Film spun cast from solution in chlorobenzene. The images show the first derivative of the actual surface heights. The cross sections of the true surface heights were taken horizontally from the points indicated by the dotted lines. Reprinted with permission from [27], C. J. Brabec et al., *Adv. Funct. Mater.* 11, 23 (2001). © 2001, Wiley-VCH.

Different strategies to improve the processability of fullerenes and/or to obtain a predetermined nanoscopic phase-separated polymer:fullerene network have been developed. For instance, systems such as diblock copolymers (conjugated donor blocks plus not conjugated fullerene-bearing blocks) [51] and conjugated oligomer–fullerene dyads have been prepared [52–60]. Peters et al. utilized an oligo(*p*–phenylene vinylene)–fulleropyrrolidine dyad to fabricate solar cells which exhibited efficiencies comparable to those of other previously reported bulk-heterojunction solar cells [59]. However, for several donor–acceptor dyads, it was found that, depending on a number of factors (e.g., aggregation state, conjugation length of the donor unit, etc.), photoinduced energy transfer can compete with photoinduced electron transfer [58–61]. Conjugated oligomer–fullerene dyads, containing donor moieties with different conjugation lengths, also have been tested as embedded components in an MDMO–PPV matrix. By comparing the performance of the obtained PV devices, it seems that dyads capable of intramolecular electron transfer (those with longer conjugated oligomer moieties) also enhance geminate carrier recombination [62, 63]. This appears plausible since, on dyads, separated charges can escape recombination only by intermolecular hopping processes. On the contrary, in systems with a long conjugated donor backbone bearing a number of acceptor fullerene moieties—such as double-cable polymers—it can be expected that holes (the positive polarons) first migrate away from electrons by a very fast intrachain diffusion process, preventing recombination. As such, double-cable polymers appear as very appealing materials that may retain the favorable electronic and photophysical properties of conjugated polymer/fullerene composites, but in which phase separation and clustering phenomena cannot occur.

4. DOUBLE-CABLE POLYMERS

A conjugated polymer/fullerene blend (a) (the occurrence of fullerene clustering is emphasized for clarity), an ideal double-cable polymer chain (b), and a more realistic picture in which the continuous pathway for the transport of electrons is also given by intermolecular contacts (c) are schematically sketched in Figure 2. Apart from aspects like the relative positioning of the polymer chains, as well as the positioning of a fullerene moiety with respect to its neighbors, a double-cable polymer for PV applications must posses the following properties:

- mutually independent ground–state electronic properties of the donor backbone and of the fullerene acceptor moieties
- as a prerequisite for photogeneration of free charge carriers, a photoinduced electron transfer from the electron-donating backbone onto the electron-accepting fullerene moieties must occur, and lead to metastable long-lived charged states.

Moreover, since the easy and cost-effective preparation of thin-film devices involves processing from solutions, solubility in common organic solvents is also a determinant factor.

A widely used route toward conjugated polymers is the electropolymerization of suitable aromatic monomers. The electropolymerization process is via electrogeneration of radical cations as reactive species and their subsequent coupling, followed by rearomatization steps [64, 65]. Several conjugated polymers, some of them carrying electroactive moieties [66], were made by synthesizing molecules specifically designed as substrates for electropolymerization [65]. Furthermore, electropolymerization allows the growth of polymeric thin films onto transparent electrodes, leading to samples suitable for most spectroscopic techniques. This is a clear advantage when the solubility of a novel conjugated polymer cannot be expected or obtained. As mentioned before, it has been reported that, in molecular donor–acceptor dyads, photoinduced energy transfer can take place in competition with intramolecular charge transfer, and geminate recombination may also be enhanced. Analougus effects, which are detrimental at least for organic PV, could occur, even in double-cable polymers (a double-cable polymer can be indeed regarded, formally, as a "polydyad"). Therefore, for the design of polymers and supramolecular structures for photovoltaic materials [67], the electrochemical approach was selected as a first step toward the preparation of double-cable polymers, as well as for the study of their electronic and photophysical properties. However, as explained above, solubility is an important requirement for practical applications. Based on the encouraging results obtained with electropolymerized insoluble double-cable polymers, the chemical synthesis of soluble and more defined double-cable nanostructures is nowadays proceeding worldwide [68].

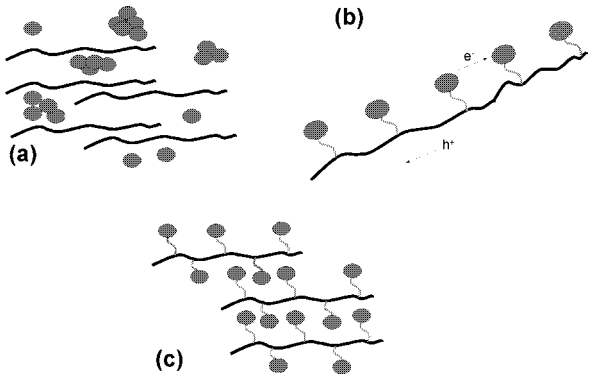
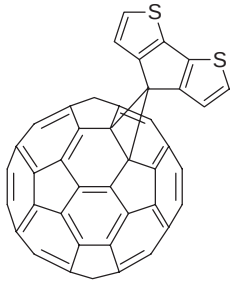


Figure 2. (a) Schematic representation of a bulk-heterojunction. Clustering of the fullerene component is emphasized. (b) An ideal, ordered double-cable polymer, and (c) a more realistic picture where interchain interactions are considered. Reprinted with permission from [68], A. Cravino and N. S. Sariciftci, *J. Mater. Chem.* 11, 1934 (2002). © 2002, Royal Society of Chemistry.

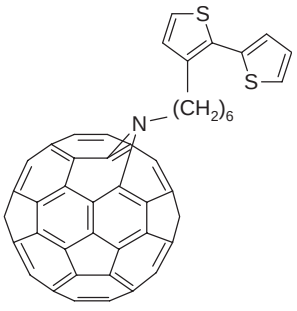
4.1. Electropolymerized Double-Cable Polymers

The first conjugated polymeric materials bearing fullerene side groups were reported by Benincori et al. and Ferraris and co-workers [63–71]. The direct grafting of C_{60} at the bridging carbon of cyclopentadithiophene, leading to compound **1** which was suitable for electropolymerization, was described in 1996 [69]. Due to its low solubility in mixtures containing acetonitrile, the anodic electropolymerization of **1** was possible only in chlorobenzene. After electrochemical compensation (dedoping), poly-**1** showed an absorption spectrum with a maximum at about 440 nm (onset above 600 nm). This λ_{max} value, which is almost 100 nm blue shifted as compared to that of electropolymerized unsubstituted cyclopentadithiophene, was explained as a consequence of the poor solubility mentioned above, leading to relatively short chains (and, therefore, to a short effective conjugation length). The electrochemical properties of poly-**1** films were studied by cyclic voltammetry in monomer-free conditions. One anodic reversible wave corresponding to the electrochemical oxidation/rereduction (*p* doping/dedoping) of the polyconjugated chain was observed at a potential 0.55 V more positive than that of the corresponding process in polycyclopentadithiophene. This result was explained as being due to the electron-withdrawing effect of the fullerene moieties. However, is worth mentioning that such an effect might originate from the short effective conjugation length as deduced from the absorption spectrum. By scanning the negative potential, three reversible waves, assigned to the multiple fullerene substituent reduction/reoxidation steps, were observed [72].



1

One year later, Ferraris et al. reported the synthesis and electrochemical polymerization of a bithiophene carrying a fullerene group through a flexible alkyl chain spacer (**2**) [70, 71]. The choice of **2** was motivated by the easier and superior electropolymerizability of bithiophenes as compared to thiophenes [64, 65]. Moreover, the flexible alkyl spacer was introduced to improve the solubility of the monomer, as well as to reduce the electron-withdrawing effect of the fullerene moiety. 2–bromo–3–(6–bromohexyl)thiophene was obtained by bromination of 3–(6–bromohexyl)thiophene with NBS in a chloroform/acetic acid solution. 3′–(6–bromohexyl)–2–2′–bithienyl was prepared by coupling the Grignard's reagent obtained from 2–bromothiophene and 2–bromo–3–(6–bromohexyl)thiophene. The reaction of the bithienyl derivative with sodium azide in DMSO afforded bythienyl azide in quantitative yeld. Treatment of the latter with C_{60} in refluxing chlorobenzene gave the monoadduct **2**, which was purified by column chromatography. In contrast to C_{60} and **1**, bithiophene **2** was indeed soluble in several organic solvents such as chloroform, dichloromethane, and THF. These improved characteristics resulted in an electropolymerized material that showed, after dedoping, an absorption maximum at $\lambda = 480$ nm. This value is comparable to that of most electrochemically synthesized polythiophenes, indicating almost no electronic or steric perturbation of the conjugated chain. Similarly, electrochemical and *in-situ* UV–vis spectroelectrochemical measurements revealed that the polymeric chain of poly-**2** behaves as common polythiophenes do. In the cathodic region of the cyclic voltammogram (CV) of poly-**2**, four fullerene reduction/reoxidation waves were observed [72]. These results indicated that, in double-cable polymers, both the donor polyconjugated chain and the fullerene acceptor moieties can mantain, to different extents, their individual electronic and electrochemical properties. In particular, by using flexible spacers of the proper length, the donor and acceptor moieties can be electronically isolated in order to prevent ground–state interactions. As such, these early works already showed that double-cable polymers could meet the first requirement for PV applications.



2

Polythiophene P3OT in combination with C_{60} or PCBM also has been utilized for the construction of promising bulk-heterojunction solar cells [73, 74], suggesting polythiophene–fullerene based double cables for further studies on PV applications. The preparation of bithiophene–fulleropyrrolidine dyad **3** was reported recently by Cravino et al. [75, 76]. The reaction of 4–hydroxybenzaldehyde with 1, 2–bis–(iodoethoxy)–ethane and K_2CO_3 afforded 4–{2–[2–(2–iodoethoxy)–ethoxy]–ethoxy}–benzaldehyde. Substitution of iodine with the potassium salt of 4–[2, 2′]bithiophenyl–4–yl–phenol gave the corresponding functionalized bithiophenyl benzaldehyde that was used in the 1, 3–dipolar azomethine ylide cycloaddition to [60]fullerene with sarcosine to afford **3** in 49% isolated yield.

Similar to the criteria already exploited by Ferraris et al. [66, 71], fulleropyrrolidine–bithiophene **3** was designed as a substrate for a relatively facile anodic electropolymerization. Both dichloromethane and toluene/acetonitrile mixtures were suitable as solvents. As a major result, the obtained poly-**3** films allowed the first observation of a photoinduced electron transfer from the polythiophene backbone to the pendant fullerene moieties within a double-cable polymer, as discussed in the following.

The cyclic voltammogram of **3** is illustrated in Figure 3. The first scan shows one irreversible wave peaking at about +1.3 V, corresponding to the oxidation of **3**. Recurrent potential scanning leads to the rising of a new redox wave around +0.8 V, corresponding to the *p* doping/dedoping process of a freshly formed poly-**3** film. Analogous results were obtained using different solvents (toluene/acetonitrile mixtures). Poly-**4** was also electrochemically prepared as reference material following the same procedure [38, 75]. The CV of poly-**3** films in a monomer-free electrolyte solution is depicted in Figure 4(a). In the positive region, one reversible process, corresponding to *p* doping/dedoping of a polythiophene backbone, was observed by the redox wave at about +0.75 V. The relationship between the maximum current peak value and the scan rate (varied from 25 to

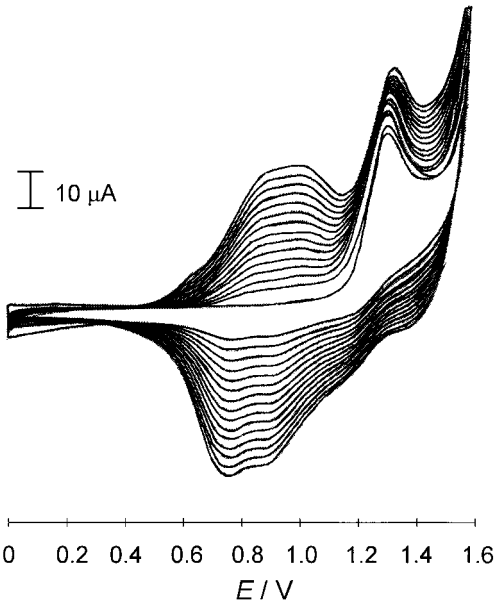
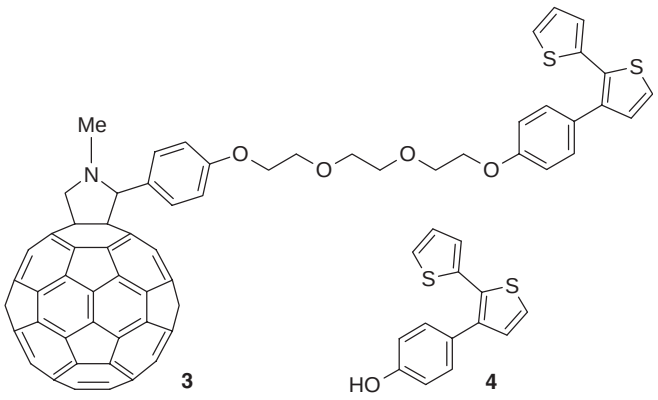


Figure 3. Anodic cyclic voltammogram of **3** (0.1 M Bu_4NPF_6in CH_2Cl_2). Working electrode: Pt foil; quasireference electrode: Ag/AgCl wire (−0.44 V versus ferrocene). Scan rate: 100 $mV \cdot s^{-1}$. Reprinted with permission from [76], A. Cravino et al., *J. Chem. Soc., Chem. Commun.* 2487 (2000). © 2000, Royal Society of Chemistry.

200 mV/s) was found to be linear [Fig. 4(b)]. A linear relationship indicates a redox-active polymer attached to the electrode, and exemplifies the stability of poly-**3** films toward *p* doping [64]. Scanning negative potentials up to −2.0 V shows several redox waves related to the multiple reduction of the fullerene moiety [72]. The irreversible peak at −0.74 V, with no clear origin, was seen only during the first cycles [Fig. 4(a)] [74]. Similar to the properties described by Ferraris et al. [70, 71], these results confirmed that both the polythiophene backbone and the pendant fullerene moieties can substantially retain their single individual electrochemical properties. Films of neutral poly-**3** were yellow–brownish, not soluble, and not photoluminescent. Their UV–vis absorption spectra showed a broad absortpion feature ranging from about 600 nm to the ITO–glass cutoff, with a pronounced shoulder at about 460 nm. As observed with poly-**1**, this value is significantly blue shifted as compared to the λ_{max} of reference poly-**4**, seen at 530 nm [75]. Again, this has been explained as an effect of the relatively low solubility of monomer **3** (as well as that of its oligomer intermediates involved in the electrochemical polymerization process), leading to a lower molecular weight for the corresponding polymer. According to the conclusions from electrochemical characterization, no hints for ground–state donor–acceptor interactions were observed.



3 4

The nature of the photoexcitations in poly-**3** solid films was investigated by means of photoinduced absorption (PIA) in the vis–NIR. The PIA spectrum, taken at an excitation wavelength of 476 nm, is shown in Figure 5. Two electronic photoinduced absorption bands are observed, one with a maximum at 1.48 eV, and one peaking below 0.6 eV. These features compare very well with PIA spectra of long oligothiophenes, and were assigned to positive polarons on a thiophene-based conjugated backbone [41, 77]. To shed light on the relaxation kinetics of the photoinduced charged states, both the intensity and modulation frequency of the pump beam were varied [78, 79]. The polaronic features, estimated by the signals at 1.38 and 0.62 eV, showed a square-root excitation intensity dependence. This dependence corresponds to a nongeminate bimolecular recombination of the photoinduced charges, suggesting the migration of charge carriers with opposite sign to different sites within the films. From the modulation frequency dependence, a broad distribution of charged state lifetimes was observed in the range from 0.8 to 10 ms.

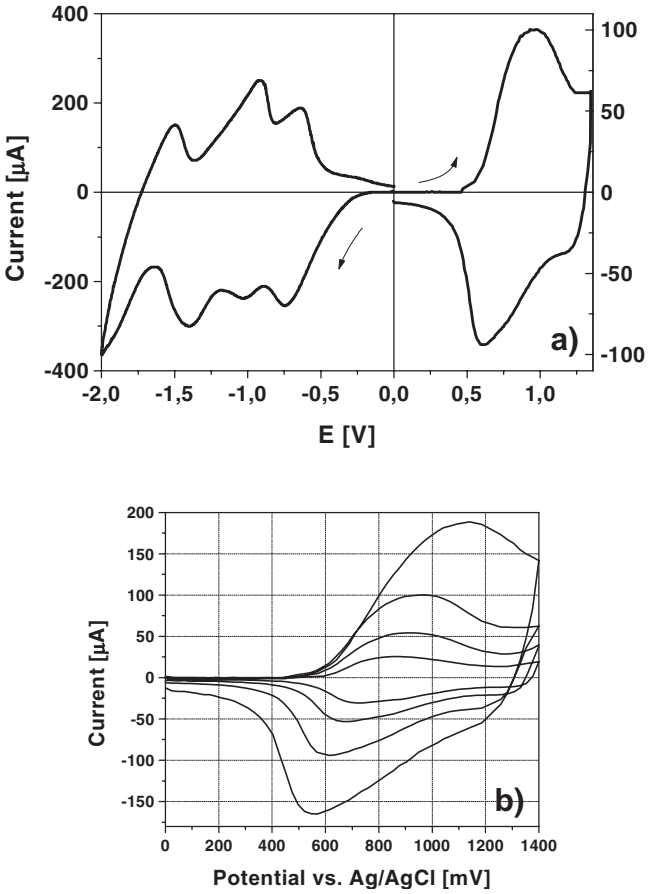


Figure 4. (a) Anodic (right) and cathodic (left) cyclic voltammogram of a poly-**3** film (0.1 M Bu_4NPF_6 in acetonitrile); conditions as in Figure 7. (b) Anodic cyclic voltammogram of poly-**3** at a scan rate of 25, 50, 100, and 200 $mV \cdot s^{-1}$. Reprinted with permission from [68], A. Cravino and N. S. Sariciftci, *J. Mater. Chem.* 11, 1934 (2002). © 2002, Royal Society of Chemistry.

While these results proved the photoinduced generation of metastable positively charged states on the polythiophene backbone, definitive evidence of a photoinduced electron transfer to the pendant fullerene moieties was obtained by

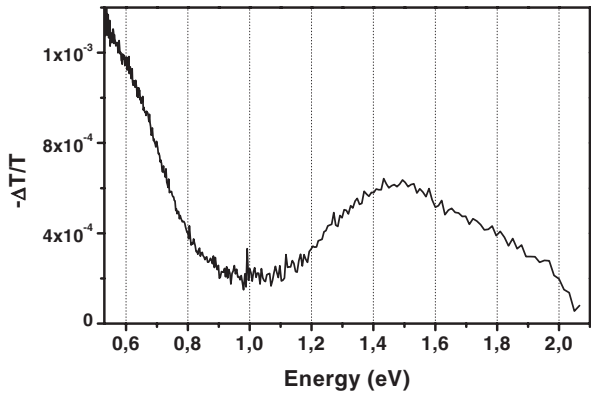


Figure 5. Photoinduced vis–NIR absorption of poly-3. Excitation at 476 nm (40 mW on a 4 mm diameter spot). $T = 100$ K. Reprinted with permission from [75], A. Cravino et al., *J. Phys. Chem. B* 106, 72 (2002). © 2002, American Chemical Society.

steady-state LESR spectroscopy. The LESR spectrum, calculated by subtracting the "light-off" signal from the "light-on" signal, shows the photogeneration of two paramagnetic species (Fig. 6). The positive polarons on the conjugated backbone were seen by the signal at a g factor of 2.0022, while the signal observed at a lower g factor, 2.0004, is typical of fullerene radical anions [43, 44]. As such, it was concluded that, in double-cable poly-3, a photoinduced electron transfer from the conjugated backbone to the pendant fulleropyrrolidine moieties leads to long-lived and mobile charge-separated states, as observed earlier in conjugated polymer/fullerene composites.

4.2. Processable Double-Cable Polymers

As already discussed, soluble double-cable polymers are required specifically for PV or other electronic applications in thin-film cost-effective devices. The attachment of fullerene moieties to conjugated backbones, which generally have low solubility, decreases the solubility even more. Therefore, the design of soluble double cables loaded with an adequate number of fullerene moieties (e.g., a fullerene concentration that allows perculation of electrons) is therefore a very challenging task for synthetic chemists. Processable double cables have been synthesized following two different strategies:

- the direct reaction of C_{60} with soluble and specifically functionalized conjugated polymers
- the copolymerization of two monomers, one carrying the fullerene moiety, and one carrying flexible chains in order to impart solubility.

The first approach has been reported by Ferraris et al. in parallel with the electrochemical preparation of poly-**2** [70]. In this work, head-to-tail copolyalkylthiophenes functionalized with azido- or primary amino groups were prepared at first. After treatment with C_{60}, soluble double cables bearing azafulleroid moieties in different percentage (**5a, b**) were obtained and isolated by extraction. The solution of these copolymers showed absorption spectra peaking at about 438 nm. This value is in rather good agreement with that of standard polythiophenes, and again, no absorption at longer wavelengths was observed. *In-situ* UV–vis absorption spectroelectrochemical measurements upon reduction of polymer **5b** showed the spectroscopical features of the fullerene moiety radical anions and, at more negative potential, of dianions, according to the results previously reported for fullerenes and azafulleroids [72, 80]. At even more negative potentials, the behavior of copolymer **5b** was similar to that reported for poly-**2**. In the anodic region, the polymer showed the typical p doping/dedoping wave of polythiophenes. As already observed with poly-**2** and poly-**3**, no hints for a ground–state interaction between the donor and the acceptor moieties were observed.

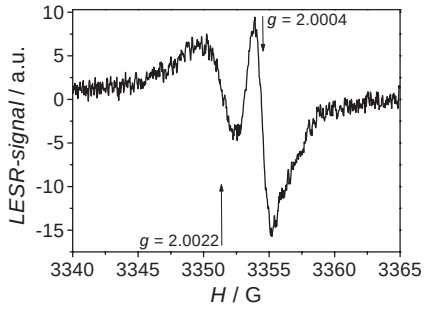
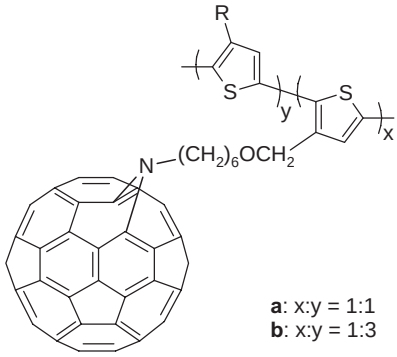


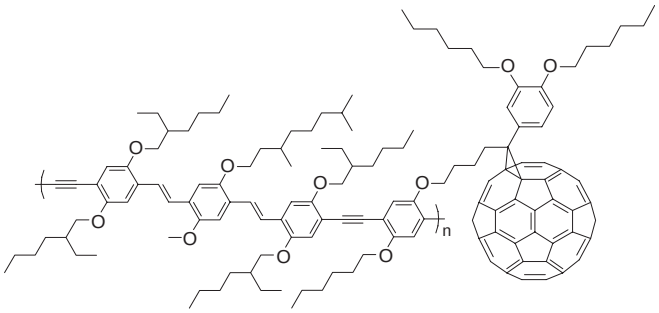
Figure 6. Integrated LESR (light-on minus light-off signal) spectrum of poly-**3** on ITO-coated plastic foils. Excitation at 488 nm (20 mW · cm^{-2}). T = 77 K. Reprinted with permission from [76], A. Cravino et al., *J. Chem. Soc., Chem. Commun.* 2487 (2000). © 2000, Royal Society of Chemistry.



5a, b

The second approach has been used recently by Marcos Ramos et al., who reported the preparation of the first PV device based on a soluble double-cable polymer [81, 82]. They synthesized a conjugated backbone hybrid of poly(p–phenylene vinylene) and poly(p–phenylene ethynylene), with covalently linked methanofullerene moieties (**6**). The polymer was synthesized under inert atmosphere in o–dichlorobenzene/triethylamine (7:3 v/v) via a palladium-catalyzed cross coupling of a diiodobenzene derivative, bearing a pendant methanofullerene, and oligo(p–phenylene vinylene) units end capped with reactive ethynylenes. A polymer similar to **6**, but lacking the methanofullerene moieties, was used as a reference for spectroscopic investigations. The effective conjugation length of the two materials was found to be nearly identical: the λ_{max} of the double-cable polymer, found at 468 nm, was only 6 nm shifted to blue as compared to that of the reference polymer. The molecular weight of **6** was M_W = 16,200 (PDI = 2.82), as estimated by means of size-exclusion chromatography. However, the authors report a behavior indicating that this value could be underestimated [80]. Spectroscopic measurements were carried out both in solution and in thin solid films. The photoluminescence of polymer **6** in a toluene solution was quenched by two orders of magnitude as compared to that of the reference polymer. The quenching of the photoluminescence in an apolar solvent probably involves a singlet–singlet energy transfer from the photoexcited conjugated backbone to the fullerene moieties, as known from studies on oligo(p–phenylene vinylene)–fulleropyrrolidine dyads [59]. However, photoluminescence quenching was also observed in thin solid films. In this case, the quenching mechanism is the photoinduced electron transfer, as discussed before. Accordingly, the photoinduced vis–NIR absorption spectrum of **6** in the solid state

exhibits a band typical of methanofullerene radical anions at about 1.2 eV (this fullerene radical-anion feature was not seen in the poly-**3** spectrum due to the concomitant presence of the broad and much more intense polaron absorption band of the polymer) accompanied by two broad bands, seen at 0.62 and 1.53 eV, characteristic of positive polarons on a conjugated chain (Fig. 7). As reported for poly-**3**, the photoinduced charge recombination was found to follow a nongeminate bimolecular mechanism. The lifetime of the photogenerated charges in **6** also extends to the millisecond time domain.



6

Photovoltaic elements were prepared by spin coating **6** from a chloroform solution [82]. The current–voltage (I/V) characteristic curves of the device recorded in the dark and under white-light illumination ($100\ \mathrm{mW \cdot cm^{-2}}$) revealed a J_{SC} value of $0.42\ \mathrm{mA \cdot cm^{-2}}$, a V_{OC} of 830 mV, and an FF of 0.29. The IPCE showed an onset at 550 nm, and a maximum of 6% at 480 nm. It is noteworthy that this performance, even though obtained with the first example of not optimized double-cable based solar cells, is already competitive with that of devices prepared following the more established p/n-heterojunction and bulk-heterojunction approaches [27, 83, 84]. Moreover, this performance was given by a double-cable polymer with a fullerene content of 31.5%-w, a value considerably smaller than that used commonly in bulk-heterojunction solar cells (about 75%-w).

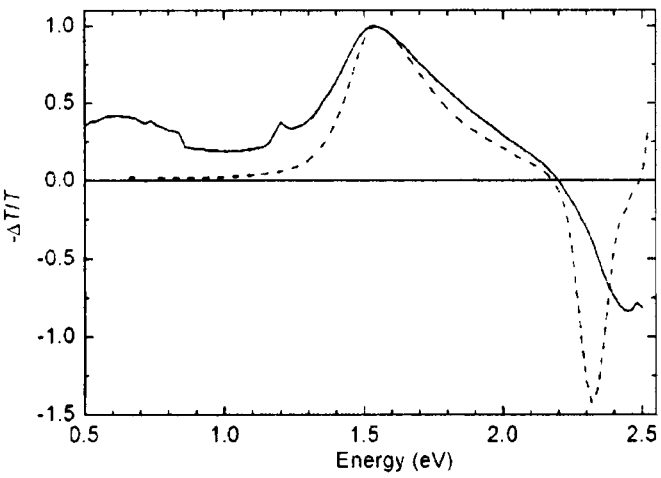
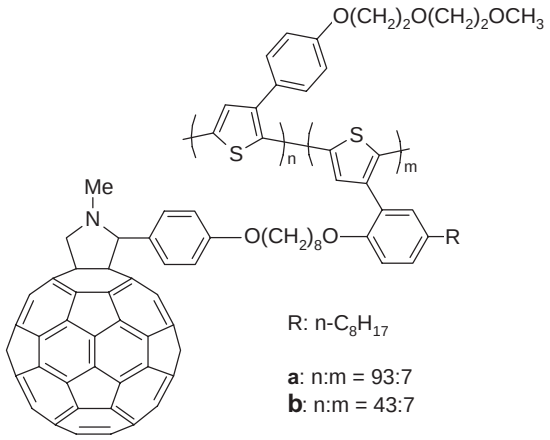


Figure 7. Photoinduced vis–NIR absorption spectra of **6** (solid line) and of the reference polymer without fullerene moieties (dashed line) on quartz. $T = 80$ K. Reprinted with permission from [81], A. Marcos Ramos et al., *J. Am. Chem. Soc.* 123, 6714 (2001). © 2001, American Chemical Society.

Another family of soluble double-cable copolythiophenes carrying fulleropyrrolidine moieties as well as solubilizing chains (**7a, b**) was reported recently by Zhang et al. [85]. These random copolymers were prepared by chemical oxidative coupling, adding a slurry of $FeCl_3$ in chloroform to the monomers dissolved in the same solvent. Two copolymers with different fullerene loadings were prepared by using two different solubilizing-monomer/fullerene-bearing monomer (prepared via the Prato reaction) ratios in the feed. After dedoping and removal of low-molecular-weight fractions and byproducts, the copolymers were found to contain about 7%-mol (corresponding to 14.5%-w, **7a**) and 14%-mol (corresponding to 24.2%-w, **7b**) of monomer-containing fullerene, respectively. Size exclusion chromatography in chloroform gave the following molecular weights: $M_W = 48{,}000$, $M_N = 28{,}000$ for **7a**, and $M_W = 39{,}000$, $M_N = 26{,}000$ for **7b**. Attempts to increase the fullerene loading were made: when higher percentages of monomer-containing fullerene were placed in the feed, an intractable material was obtained. When the solubilizing monomer used in excess in this preparation is used alone, a polythiophene known as PEOPT is obtained (see Chart 1). PEOPT spin coated from a chloroform solution leads to an orange phase. This phase can be converted to a more ordered blue phase with extended effective conjugation by thermal annealing or by chloroform vapor exposure [46, 86]. It is of interest to note that **7a** and **7b** retain this property, despite the presence of the bulky fullerene moieties. Both copolymers undergo a color change from orange to blue upon exposure to chloroform vapors at room temperature. In this blue phase, both copolymers display a broad electronic absorption band covering a wide range of the solar spectrum, which is in itself an interesting property for PV materials.



7a, b

The solubility of **7a** and **7b** allowed the preparation of photodiodes [85]. The devices based on the double cable with the higher fullerene loading **7b** showed an IPCE more than twofold higher as compared to that of the devices prepared with **7a** (Fig. 8). In both cases, a significant

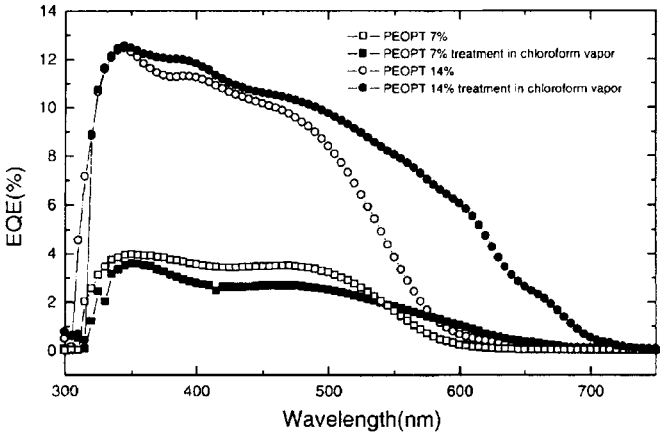
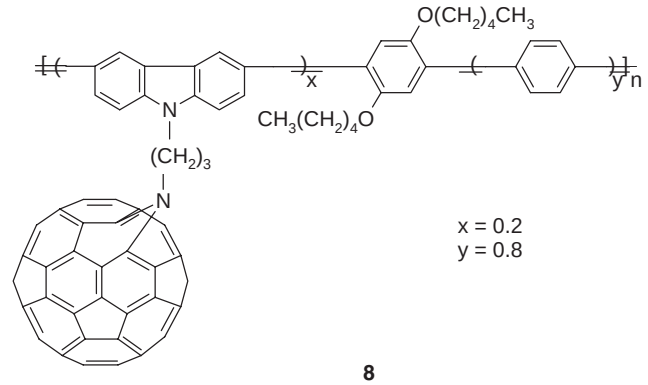


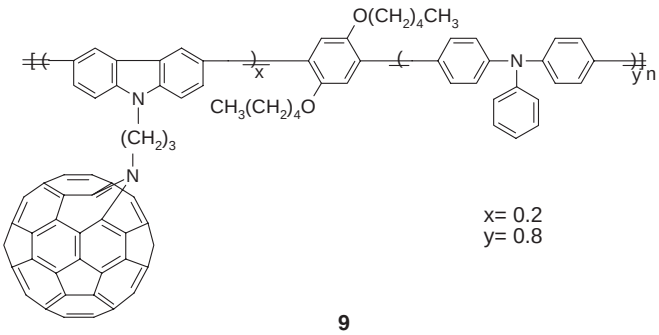
Figure 8. External quantum efficiency (IPCE) of ITO/PEDOT:PSS/**7a** or **7b**/Al photodiodes under monochromatic light illumination prior to and after conversion from the orange to the blue phase. Reprinted with permission from [85], F. Zhang et al., *Adv. Funct. Mater.* 13, 171 (2001). © 2001, Wiley-VCH.

improvement of the IPCE was obtained after treatment of the photoactive films with chloroform just before evaporation of the Al electrode. This effect reflects the change in the absorption spectra while going from the orange to the blue phase of the copolymers. The devices prepared using **7b** where characterized under monochromatic light (λ = 505 nm) illumination with an intensity of 0.1 $mW \cdot cm^{-2}$. The FF was 0.25, and the power conversion efficiency was 0.6% [49]. The diodes not treated with chloroform vapor gave the highest photovoltages, V_{OC} = 750 mV. This value is higher than that reported for PEOPT/C_{60} bulk-heterojunction solar cells [46]. After treatment with chloroform vapors, the V_{OC} of the device dropped to 450 mV. This behavior is a consequence of the altered energy levels between the orange and the blue phases [37, 45, 85].

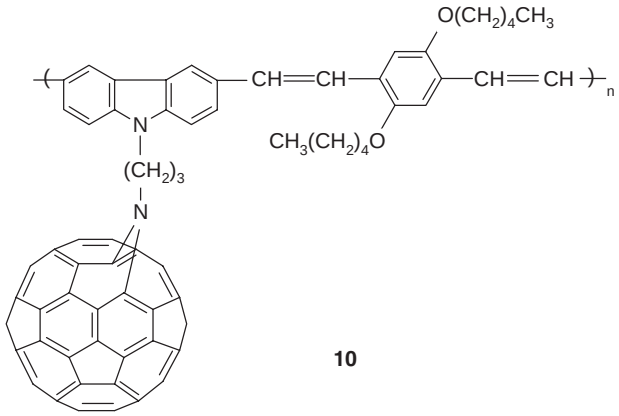
Another class of soluble double cables was described by Xiao and by Wang et al. [87, 88]. Based on the well-known photoconducting as well as electrochemical properties of carbazole, they synthesized a soluble regular copolymer containing *p*–phenylene vinylene and carbazole units, the latter bearing a pendant azido functionality. The reaction with C_{60} in chlorobenzene gave a soluble fraction (CCl_4 as solvent) with a fullerene percentage estimated spectroscopically as 19.9%-w (**8**). Similarly, they synthesized a double cable also containing triphenylamine moieties (**9**). In this case, the fullerene content was about 21%-w. The weight-average M_w values were determined by means of gel permeation chromatography as 3810 for **8** and 5051 for **9**. Based on these limited values, it is clear that these double-cable materials consist of only a few repeating units. Highly fullerene-loaded double cable **10** was prepared by the same group [88]. In this latter material, the fullerene percentage was estimated as high as 68.8%-w. This value seems to be overestimated since, based on the repeating unit, the calculated maximum fullerene loading is about 57%-w. Nevertheless, this is a high fullerene loading, approaching those of conjugated polymer/fullerene composites used for the fabrication of bulk-heterojunction solar cells with improved efficiency. By means of thermal gravimetric analysis, these copolymers have been found to be stable up to about 350 °C, and moderate weight losses occurred just within the temperature range 380–460 °C. Spectroscopic and electrochemical studies allowed the authors to conclude that the electronic properties of the fullerene moieties were substantially preserved in all cases. Photoinduced electron transfer has been claimed to occur due to the photoluminescence quenching observed in chloroform solutions. These double-cable polymers are currently being tested as material for plastic PV energy conversion [88].



8



9

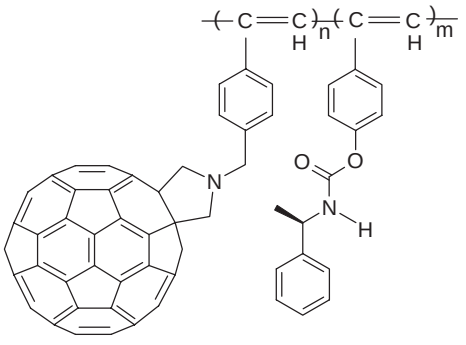


10

4.2.1. An Optically Active Double-Cable *Polymer*

Nishimura et al. synthesized and characterized optically active helical polyphenylacetylenes carrying tethered achiral fullerene groups (**11**) [89]. The latter arrange in

a helical array with a predominant screw sense along the conjugated backbone. Optically active *cis-transoidal* polyphenylacetylene double cables **11** were prepared by rhodium-catalyzed copolymerization of an achiral fulleropyrrolidine phenylacetylene, synthesized using Prato's reaction, and an optically active phenylacetylene with an (*R*)–[(1–phenylethyl)carbamoyil]oxy unit on the phenyl group. The fullerene homopolymer was insoluble, whereas the copolymers with a fullerene unit content of 10, 20, and 30%-mol were partially soluble in THF or chloroform. The circular dichroism (CD) and absorption spectra of the copolymer containing 10%-mol of fullerene monomer are shown in Figure 9. This copolymer exhibited an intense induced CD in the region where the main conjugated backbone absorbs, the magnitude of which increases monotonically with decreasing temperature. The same effect also can be observed with the homopolymer containing no fullerene monomer. In dilute solutions, the copolymer does not exhibit detectable absorption and CD above 500 nm due to the low fullerene content. At higher concentrations and at a temperature below −40 °C, remarkable CD induction was observed in the characteristic region of the fullerene chromophore (over 700 nm), and the induced CD intensity further increased at −80 °C (see inset in Fig. 9). This behavior indicates that the achiral fullerene units of the copolymer arrange in a helical array with a one-handed screw sense. Similar results were found with the copolymers containing higher fullerene percentages. This seems to be the first example of an induced chirality upon achiral fullerene units, derived from their induced helical arrangment *along a conjugated backbone*. As such, this work demonstrates that the proper design of the double-cable primary structure allows the attainment of a defined nanostructure, both in terms of distance between the acceptor units and the conjugated polymer skeleton, as well as of the positioning of the single acceptor moieties with respect to their neighbors. In fact, the authors have proposed a helical structure (Fig. 10) which suggests similar approaches as a way to improve the intercontact between fullerene moieties of a double-cable polymer, even with relatively low fullerene loading.



11

5. ALTERNATIVES TO FULLERENE-CONTAINING POLYMERS

Fullerene derivatives so far have been the most promising electron acceptor materials for donor–acceptor polymeric solar cells, as well as electron-transporting components

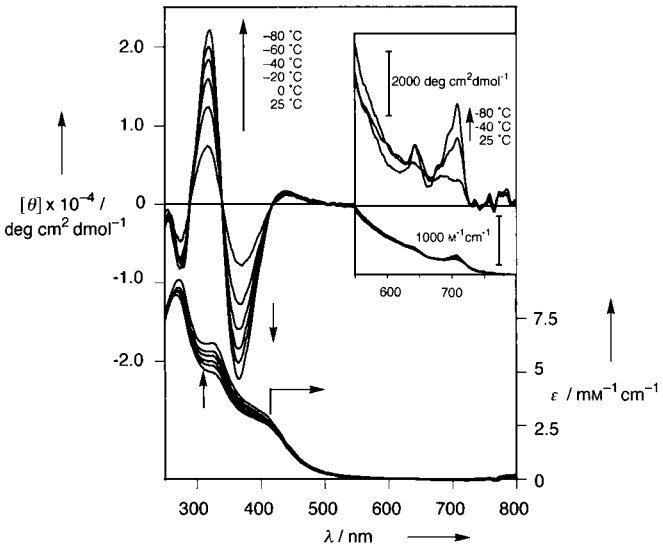


Figure 9. CD and absorption spectra of poly-**11** with a fullerene fraction of 20%-mol in THF solution (0.11 mg · mL^{-1}) at various temperatures. The inset shows the corresponding CD and absorption spectra in the fullerene region in THF at a higher concentration (4.6 mg · mL^{-1}). The molar concentrations were calculated based on the monomer units and C_{60} units (inset). Reprinted with permission from [89], T. Nishimura et al., *Angew. Chem.* 114, 3754 (2002). © 2002, Wiley-VCH.

in organic electronic devices. However, a vast range of electron acceptor molecules are known, which are readily available and could be preferred species to fullerenes, especially from the preparative, cost-effective, and tunable viewpoints. Zerza et al. prepared and characterized a double-cable polymer where the electron acceptor moieties were of the tetracyanoanthraquinone type [35]. This polythiophene-based double-cable polymer was synthesized electrochemically, starting from the bithiophene derivative **12**. As observed for fullerene-containing double-cable polymers, the electrochemical and electronic properties of the individual conjugated chain and of the tethered tetracynoanthraquinone-type moieties in poly-**12** were nicely preserved [35]. Again, a photoinduced electron transfer from the polythiophene chain to the acceptor moieties, leading to long-lived charged states, was demonstrated by comparing electrochemical and photoexcited spectroscopy results.

Finally, a quite different approach was proposed by Skabara et al. [36]. The authors reported the chemical and electrochemical preparation of a donor–acceptor noncomposite polymeric material consisting of a polythiophene incorporating π-conjugated 1, 3–dithiole–2–ylidenefluorene (**13**). Different from the materials described above, the donor and the acceptor components (the polythiophene chain and the 1, 3–dithiole–2–yledenefluorene fragments, respectively) of this nanostructure are connected through sulphur bridges which allow electronic interactions between them. Indeed, internal charge transfer occurs in the monomer as well as in **13**, as seen by electronic absorption measurements. However, despite the occurrence of this ground–state electronic interaction, polymer **13** undergoes photoinduced charge separation, as seen by photoinduced IR absorption [36].

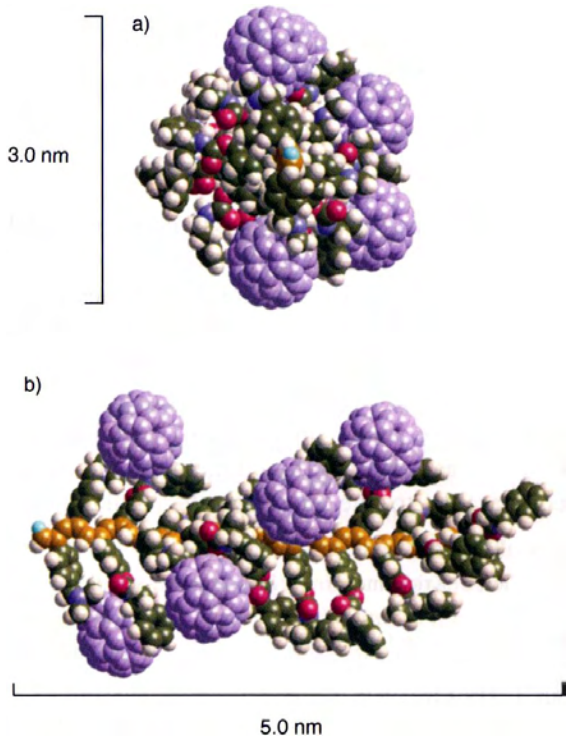
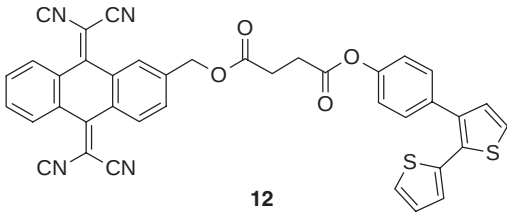
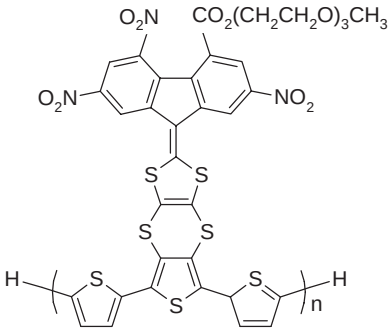


Figure 10. Possible helical structure of poly-**11** with a fullerene fraction of 20%-mol (20-mer). Space-filling models are shown in the top (a) and side views (b). The fullerene units are assumed to arrange in a left-handed helical array along the right-handed polyacetylene main chain. The helix sense is tentative, but the copolymer should have a helical conformation (either a right- or left-handed helix), as it exhibits an ICD in both the main chain and the fullerene chromophoric regions. Reprinted with permission from [89], T. Nishimura et al., *Angew. Chem.* 114, 3754 (2002). © 2002, Wiley-VCH.



12



13

6. SUMMARY

The preparation and properties of a novel class of fullerene functional materials such as double-cable polymers have been reviewed. These nanostructures consist of a hole-conducting, conjugated backbone (*p* cable) with covalently bound fullerene moieties (*n* cable). By the choice of the proper covalent linkage, both in terms of chemical nature and length, the donor and the acceptor moieties can be electronically "isolated" to exclude ground–state interactions. As such, double-cable polymers are attractive for all organic electronic applications in which bipolar transport is required. Photoexcitation spectroscopic investigations have shown that double cables in their solid state undergo photoinduced electron transfer, leading to long-lived, mobile charge carriers, as observed earlier in conjugated polymer/fullerene composites. Since, in double-cable polymers, phase separation cannot occur, these materials are indeed appealing as a way to control both electronic and nanostructural properties within the photoactive layer of plastic solar cells. The very important technological issue is processability, which is certainly the major challenge toward practical application of double-cable polymers in thin-film devices. With the few soluble double-cable polymers prepared so far, the fabrication of prototype photodiodes and solar cells has been demonstrated. However, so far, a compromise between the fullerene loading (for the transport of electrons) and solubility has been made. In this respect, the following hold.

- Improving the chemical design of a double cable may allow the preparation of processable materials with an acceptably high fullerene moieties percentage as well as acceptable solubility.
- Fullerene content as high as 14%-mol (24.2%-w) has been obtained in a soluble random copolymer (**7b**), and a loading of 31.5%-w already has been obtained, even in a polymer with a well-defined backbone (**7**).
- Balberg et al. described bipolar transport in a poly(3–hexylthiophene)/C_{60} composite with a fullerene fraction of 10%-mol [90]. Indeed, it should be taken into account that the volume fraction occupied by the fullerene moieties could be enhanced by minimizing the volume occupied by the solubilizing but electrically inactive side chains.
- The double-cable concept may lead to bipolar transport properties, even in guest–host systems where the double-cable polymer is embedded within processable polymers as matrices.
- Synthetic strategies toward fullerene derivatives able to self-assemble in well-defined nanostructures with suitably designed conjugated polymers should be considered as an alternative to covalently bound double cables, too. To this end, self-assembly strategies based on electrostatically driven or other key/lock interactions will be possible [91–93].

In addition, to gain additional control of the final morphology by their tertiary structure, covalently linked double-cable polymers also could be further functionalized [91–94].

GLOSSARY

AM1.5 illumination conditions Sunlight radiating through an air mass 1.5 times greater than the vertical case (typical standard definition for specific solar radiation conditions).

Bipolar organic polymer An organic polymer able to stabilize and transport both positive and negative charge carriers.

Conjugated polymer A polymer whose backbone consists of unsaturated carbon atoms or of the repetition of unsaturated units which may allow π-electron delocalization. Well known conjugated polymers are, for instance, polyacetylene $(—CH=CH—)_n$ and polythiophene $(—Th—)_n$.

Doping of conjugated polymers Term developed in analogy with the terminology used in the field of inorganic semiconductors. In conjugated polymers, and more generally in organic semiconductors, doping is a chemical or electrochemical process (mostly of redox type) upon which charged species (for instance polarons) are generated. As result the conductivity of the material is increased. Since photoexcitation may lead to substantially similar results, the term photodoping is also used often.

Double-cable polymer A bipolar conjugated polymer consisting of a conjugated backbone, which acts as electron donor and transporting moiety ("*p*-cable"), bearing tethered or grafted substituents that act as electron acceptor and transporting moieties ("*n*-cable").

Fill factor (of a photovoltaic element) The fill factor (FF) of a solar cell is determined as $FF = I_{mpp}V_{mpp}/I_{SC}V_{OC}$, where I_{mpp} and V_{mpp} are the current and voltage at the maximum power point in the fourth quadrant of the I/V curve and I_{SC}, V_{OC} are the short circuit current and the open circuit voltage, respectively.

Incident photons to converted electrons (IPCE) Ratio of the number of electrons collected at the electrodes versus the number of incident photons; also called external quantum efficiency (EQE).

Photoinduced electron transfer A photoinduced redox reaction between an electron donor, which acts as reducing species, and an electron acceptor, which acts as oxidizer. After photoinduced electron transfer the two species are positively and negatively charged.

Polarons Radical-ions on systems with extended π-electron delocalization. Due to the strong electron-phonon coupling in these systems, an electronic perturbation is usually accompanied by a considerable distortion of the molecular geometry.

Polymer:fullerene plastic solar cells Photovoltaic devices constituted of a conjugated polymer and a fullerene derivative sandwiched between electrodes. After photoinduced electron transfer from the conjugated polymer onto the fullerene, the separated charge carriers are collected at the electrodes and used to deliver power to an external utilizer. These solar cells can be fabricated by well developed coating techniques (e.g. spin coating from solution) on flexible and light weight plastic substrates coated with conductive materials. Thus, they are often called plastic solar cells.

Power efficiency of a photovoltaic element The power-in to power-out conversion efficiency η is given by $\eta = I_{SC} \cdot V_{OC}FF/I_{light}$, where I_{light} is the power of the incident light.

ACKNOWLEDGMENTS

Michele Maggini is gratefully aknowledged for helpful discussions.

REFERENCES

1. H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl, and R. E. Smalley, *Nature* 318, 162 (1985).
2. W. Krätschmer, L. D. Lamb, K. Fostiropoulos, and D. R. Hufman, *Nature* 347, 354 (1990).
3. M. S. Dresselhaus, G. Dresselhaus, and P. C. Eklund, "Science of Fullerenes and Carbon Nanotubes." Academic, San Diego, 1996.
4. F. Diederich and C. Thilgen, *Science* 271, 317 (1996).
5. A. Hirsch, "Fullerenes and Related Structures, Topics in Current Chemistry," Vol. 199. Springer-Verlag, Berlin, 1999.
6. J. L. Segura and N. Martín, *J. Mater. Chem.* 2403 (2000) and references therein.
7. M. Prato, *J. Mater. Chem.* 7, 1907 (1997) and references therein.
8. K. Kadish and R. Ruoff (Eds.), "Fullerenes: Chemistry, Physics and Technology." Wiley Interscience, New York, 2000.
9. A. J. Heeger, *J. Phys. Chem. B* 105, 8475 (2001).
10. H. S. Nalwa (Ed.), "Handbook of Organic Conductive Molecules and Polymers," Vols. 1–4. Wiley, Chichester, 1997.
11. T. A. Skotheim, R. L. Elsenbaumer, and J. R. Reynolds (Eds.), "Handbook of Conducting Polymers," 2nd ed. Marcel Dekker, New York, 1998.
12. G. Hadziioannu and P. F. van Hutten (Eds.), "Semiconducting Polymers, Chemistry Physics and Engineering." Wiley, Weinheim, 2000.
13. J. H. Burroughs, D. D. C. Bradley, A. B. Brown, R. N. Marks, K. Mackay, R. H. Friend, P. L. Burn, and A. B. Holmes, *Nature* (*London*) 347, 539 (1990).
14. G. Gustafsson, Y. Cao, G. M. Treacy, F. Klavetter, N. Colaneri, and A. J. Heeger, *Nature* (*London*) 357, 477 (1992).
15. G. Grem, G. Leditzky, B. Ulrich, and G. Leising, *Adv. Mater.* 4, 36 (1992).
16. A. Kraft, A. C. Grimsdale, and A. B. Holmes, *Angew. Chem., Int. Ed. Engl.* 37, 402 (1998).
17. N. S. Sariciftci, L. Smilowitz, A. J. Heeger, and F. Wudl, *Science* 258, 1474 (1992).
18. S. Morita, A. A. Zakhidov, and K. Yoshino, *Solid State Commun.* 82, 249 (1992).
19. N. S. Sariciftci, D. Braun, C. Zhang, V. I. Srdanov, A. J. Heeger, G. Stucky, and F. Wudl, *Appl. Phys. Lett.* 62, 585 (1993).
20. G. Yu, J. Wang, J. McElvain, and A. J. Heeger, *Adv. Mater.* 10, 1431 (1998).
21. H. Meier, "Organic Semiconductors." Verlag-Chemie, Weinheim, 1974.
22. C. W. Tang, *Appl. Phys. Lett.* 48, 183 (1986).
23. J. Simon and J.-J. André, "Molecular Semiconductors." Springer, Berlin, 1985.
24. M. Kaneko, in "Handbook of Organic Conductive Molecules and Polymers" (H. S. Nalwa, Ed.), Vol. 4, p. 661. Wiley, Chichester, 1997.
25. J. Rostalski and D. Meissner, *Sol. Energy Mater. Sol. Cells* 61, 87 (2000).
26. T. Tsuzuki, Y. Shirota, J. Rostalski, and D. Meissner, *Sol. Energy Mater. Sol. Cells* 61, 1 (2000).
27. C. J. Brabec, N. S. Sariciftci, and J. C. Hummelen, *Adv. Funct. Mater.* 11, 15 (2001) and references therein.
28. G. Yu, J. Gao, J. C. Hummelen, F. Wudl, and A. J. Heeger, *Science* 270, 1789 (1995).
29. S. E. Shaheen, C. J. Brabec, F. Padinger, T. Fromherz, J. C. Hummelen, and N. S. Sariciftci, *Appl. Phys. Lett.* 78, 841 (2001).
30. Y. Wang and A. Suna, *J. Phys. Chem. B* 101, 5267 (1997).
31. J. J. M. Halls, C. A. Walsh, N. C. Greenham, E. A. Marseglia, R. H. Friend, S. C. Moratti, and A. B. Holmes, *Nature* 270, 1789 (1995).

32. G. Yu and A. J. Heeger, *J. Appl. Phys.* 78, 4510 (1995).
33. G. Zerza, M. C. Scharber, C. J. Brabec, N. S. Sariciftci, R. Gómez, J. L. Segura, N. Martín, and V. I. Srdanov, *J. Phys. Chem. A* 104, 8315 (2000).
34. T. Yohannes, H. Neugebauer, S. A. Jenekhe, and N. S. Sariciftci, *Synth. Met.* 116, 241 (2001).
35. G. Zerza, A. Cravino, H. Neugebauer, N. S. Sariciftci, R. Gómez, J. L. Segura, N. Martín, M. Svensson, and M. R. Andersson, *J. Phys. Chem. A* 105, 4172 (2001).
36. P. J. Skabara, I. Serebrykov, I. Perepichka, N. S. Sariciftci, H. Neugebauer, and A. Cravino, *Macromolecules* 34, 2232 (2001).
37. C. J. Brabec, A. Cravino, D. Meissner, N. S. Sariciftci, T. Fromherz, M. T. Rispens, L. Sanchez, and J. C. Hummelen, *Adv. Funct. Mater.* 11, 374 (2001).
38. H. Neugebauer, C. J. Brabec, J. C. Hummelen, and N. S. Sariciftci, *Sol. Energy Mater. Sol. Cells* 61, 35 (2000).
39. N. S. Sariciftci (Ed.), "Primary Photoexcitations in Conjugated Polymers: Molecular Exciton versus Semiconductor Band Model." World Scientific, Singapore, 1997.
40. R. Menon, C. O. Yoon, D. Moses, and A. J. Heeger, in "Handbook of Organic Conductive Molecules and Polymers" (H. S. Nalwa, Ed.), Vol. 1, Chap. 2, and references therein. Wiley, Chichester, 1997.
41. L. Smilowitz, N. S. Sariciftci, R. Wu, C. Gettinger, A. J. Heeger, and F. Wudl, *J. Phys. Chem. B* 47, 13835 (1993).
42. C. J. Brabec, G. Zerza, G. Cerullo, S. De Silvestri, S. Luzzati, J. C. Hummelen, and N. S. Sariciftci, *Chem. Phys. Lett.* 340, 232 (2001).
43. P. M. Allemand, G. Srdanov, A. Koch, K. Khemeni, F. Wudl, Y. Rubin, F. N. Diederich, M. M. Alvarez, S. J. Anz, and R. L. Whetten, *J. Am. Chem. Soc.* 113, 2780 (1991).
44. V. Dyakonov, G. Zoriniants, M. C. Scharber, C. J. Brabec, R. A. J. Janssen, J. C. Hummelen, and N. S. Sariciftci, *Phys. Rev. B* 59, 8019 (1999).
45. C. J. Brabec, A. Cravino, D. Meissner, N. S. Sariciftci, M. T. Rispens, L. Sanchez, J. C. Hummelen, and T. Fromherz, *Thin Solid Films* 403–404, 368 (2002).
46. L. S. Roman, W. Mammo, L. A. A. Petersson, M. R. Andersson, and O. Inganäs, *Adv. Mater.* 10, 774 (1998).
47. J. J. M. Halls and R. H. Friend, *Synth. Met.* 85, 1307 (1997).
48. T. Martens, J. D'Haen, T. Munters, L. Goris, Z. Beelen, J. Manca, M. D'Olieslaeger, D. Vanderzande, L. De Schepper, and R. Andriessen, *Mater. Res. Soc. Symp. Proc.* 725, P7.11 (2002).
49. The fill factor (FF) of a photovoltaic element is determined as $FF = I_{mpp}V_{mpp}/I_{SC}V_{OC}$, where I_{mpp} and V_{mpp} are the current and voltage at the maximum power point in the fourth quadrant of the I/V curve. The power-in to power-out conversion efficiency η is given by $\eta = I_{SC}V_{OC}FF/I_{light}$, where I_{light} is the intensity of the incident light. AM1.5 illumination conditions correspond to the global sun light spectrum, with $I_{light} = 860\ W \cdot m^{-2}$. An interesting discussion on the usefulness of different kinds of given efficiency values can be found in [25].
50. W. Geens, J. Poortmans, S. E. Shaeen, C. J. Brabec, N. S. Sariciftci, and B. Wessling, *Organ. Electron.* 3, 105 (2002).
51. U. Stalmach, B. de Boer, C. Videlot, P. F. van Hutten, and G. Hadzioannou, *J. Am. Chem. Soc.* 122, 5464 (2000).
52. N. Martín, L. Sanchez, B. Mescas, and I. Pérez, *Chem. Rev.* 98, 2527 (1998).
53. S.-G. Liu, L. Shu, J. Rivera, H. Liu, J.-M. Raimundo, J. Roncali, A. Gorgues, and L. Echegoyen, *J. Organ. Chem.* 64, 4884 (1999).
54. S. Knorr, A. Grupp, M. Mehring, G. Grube, and F. Effenberger, *J. Chem. Phys. A* 110, 3502 (1999).
55. T. Yamashiro, Y. Aso, T. Otsubo, H. Tang, Y. Harima, and K. Yamashita, *Chem. Lett.* 443 (1999).
56. J.-F. Nierengarten, J.-F. Eckert, J.-F. Nicoud, L. Ouali, V. V. Krasnikov, and G. Hadziioannou, *J. Chem. Soc. Chem. Commun.* 617 (1999).
57. P. A. van Hal, J. Knol, B. M. W. Langeveld-Voss, S. C. J. Meskers, J. C. Hummelen, and R. A. J. Janssen, *J. Phys. Chem. A* 104, 5974 (2000).
58. N. Armaroli, F. Barigelletti, P. Ceroni, J.-F. Eckert, J.-F. Nicoud, and J.-F. Nierengarten, *J. Chem. Soc., Chem. Commun.* 599 (2000).
59. E. Peters, P. A. van Hal, J. Knol, C. J. Brabec, N. S. Sariciftci, J. C. Hummelen, and R. A. J. Janssen, *J. Phys. Chem. B* 104, 10174 (2000).
60. J. L. Segura, R. Gómez, N. Martín, C. Luo, and D. M. Guldi, *J. Chem., Soc., Chem. Commun.* 701 (2000).
61. I. B. Martini, B. Ma, T. Da Ros, R. Helgeson, F. Wudl, and B. Schwartz, *Chem. Phys. Lett.* 327, 253 (2000).
62. A. Cravino, C. J. Brabec, A. Y. Andreev, R. Rittberger, N. S. Sariciftci, E. Peeters, R. A. J. Janssen, J. Knol, and J. C. Hummelen, "MRS 2000 Fall Meeting, Cluster 8: Organic and Biomaterials, Symposium JJ: Organic Electronic and Photonic Materials and Devices," Boston, 2000.
63. A. Cravino, Ph.D. Dissertation, Linz, 2002.
64. J. Simonet and J. Rault Berthelot, *Prog. Solid State Chem.* 21, 1 (1991).
65. J. Roncali, *J. Mater. Chem.* 9, 1875 (1999) and references therein.
66. J. Roncali, *J. Mater. Chem.* 9, 1888 (1999).
67. European Commision, Joule III Program for Molecular Plastic Solar Cells, Contract No. JOR3CT980206.
68. A. Cravino and N. S. Sariciftci, *J. Mater. Chem.* 12, 1931 (2002).
69. T. Benincori, E. Brenna, F. Sannicoló, L. Trimarco, and G. Zotti, *Angew. Chem.* 108, 718 (1996).
70. J. P. Ferraris, A. Yassar, D. Loveday, and M. Hmyene, *Opt. Mater.* (*Amsterdam*) 9, 34 (1998).
71. A. Yassar, M. Hmyene, D. C. Loveday, and J. P. Ferraris, *Synth. Met.* 84, 231 (1997).
72. L. Echegoyen and L. E. Echegoyen, *Acc. Chem. Res.* 31, 593 (1998).
73. D. Gebeheyu, F. Padinger, T. Fromherz, J. C. Hummelen, and N. S. Sariciftci, *Int. J. Photoenergy* 1, 95 (1999).
74. D. Gebeheyu, F. Padinger, T. Fromherz, J. C. Hummelen, and N. S. Sariciftci, *Bull. Chem. Soc. Ethiop.* 14, 57 (2000).
75. A. Cravino, G. Zerza, H. Neugebauer, M. Maggini, S. Bucella, E. Menna, M. Svensson, M. R. Andersson, C. J. Brabec, and N. S. Sariciftci, *J. Phys. Chem. B* 106, 70 (2002).
76. A. Cravino, G. Zerza, M. Maggini, S. Bucella, M. Svensson, M. R. Andersson, H. Neugebauer, and N. S. Sariciftci, *J. Chem. Soc., Chem. Commun.* 2487 (2000).
77. R. A. J. Janssen, D. Moses, and N. S. Sariciftci, *J. Chem. Phys.* 101, 9519 (1996).
78. G. Dellepiane, C. Cuniberti, D. Comoretto, G. F. Musso, G. Figari, A. Piaggi, and A. Borghesi, *Phys. Rev. B* 48, 7850 (1993).
79. C. Botta, S. Luzzati, R. Tubino, D. D. C. Bradley, and R. H. Friend, *Phys. Rev. B* 48, 14809 (1993).
80. M. K. Mazen, R. T. Carlin, P. C. Trulove, G. R. Eaton, and S. S. Eaton, *J. Am. Chem. Soc.* 116, 3465 (1994).
81. A. Marcos Ramos, M. T. Rispens, J. K. J. van Duren, J. C. Hummelen, and R. A. J. Janssen, *J. Am. Chem. Soc.* 123, 6714 (2001).
82. A. Marcos Ramos, M. T. Rispens, J. C. Hummelen, and R. A. J. Janssen, *Synth. Met.* 119, 171 (2001).
83. L. S. Roman, M. R. Andersson, T. Yohannes, and O. Inganäs, *Adv. Mater.* 9, 1164 (1997).
84. H. Matoussi, M. F. Rubner, F. Zhou, J. Kumar, S. K. Tripathy, and L. Y. Chiang, *Appl. Phys. Lett.* 77, 1540 (2000).
85. F. Zhang, M. Svensson, M. R. Andersson, M. Maggini, S. Bucella, E. Menna, and O. Inganäs, *Adv. Mater.* 13, 171 (2001).
86. C. J. Brabec, C. Winder, M. C. Scharber, N. S. Sariciftci, J. C. Hummelen, M. Svensson, and M. R. Andersson, *J. Chem. Phys.* 115, 7235 (2001).

87. S. Xiao, S. Wang, H. Fang, Y. Li, Z. Shi, C. Du, and D. Zhu, *Macromol. Rapid. Commun.* 22, 1313 (2001).
88. S. Wang, S. Xiao, Y. Li, Z. Shi, C. Du, H. Fang, and D. Zhu, *Polymer* 43, 2049 (2002).
89. T. Nishimura, K. Takatani, S. Sakurai, K. Maeda, and E. Yashima, *Angew. Chem.* 114, 3754 (2002).
90. I. Balberg, R. Naidis, M.-K. Lee, J. Shinar, and L. F. Fonseca, *Appl. Phys. Lett.* 79, 197 (2001).
91. M. T. Rispens, L. Sánchez, J. Knool, and J. C. Hummelen, *J. Chem. Soc., Chem. Commun.* 161 (2001).
92. J. J. González, S. González, E. M. Prego, C. Luo, D. M. Guldi, J. de Mendoza, and N. Martín, *J. Chem. Soc., Chem. Commun.* 163 (2001).
93. S. Mizyed, P. E. Georghiou, M. Barcu, B. Cuadra, A. K. Rai, P. Cheng, and L. T. Scott, *J. Am. Chem. Soc.* 123, 12770 (2001).
94. S. A. Jenekhe and X. L. Chen, *Science* 279, 1903 (1998).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanostructured Chalcogenide Glasses

Keiji Tanaka

Hokkaido University, Sapporo, Japan

CONTENTS

1. INTRODUCTION

Nanostructures can be defined as having a typical length scale between 1 nm and a few 100 nm. As shown in Figure 1, this scale covers molecular and cluster dimensions, and it is clearly greater than the atomic distance of approximately 0.2 nm. On the other hand, this scale is substantially smaller than the submicrometer wavelength of visible light, which can be regarded as a macroscopic scale. The nanostructure may also be defined as the structure in which material properties depend on the sample size. For instance, in a cubic solid with a side length of 10 nm, the number n_s of atoms located at the surface is nearly equal to that n_b included in the interior, and the ratio n_s/n_b dramatically increases with a decrease in the side length. In such small solids, the surface, not the bulk, may govern material properties.

For amorphous materials, the control of nanoscale structures could be an ultimate research subject. This situation is in contrast to that of single crystalline materials, in which *atomic* structures can be controlled. For instance, we can prepare atomically flat surfaces, where atoms can be controlled one by one. In contrast, in amorphous materials such atomic manipulations are meaningless, since the amorphous structure is disordered at the atomic level. Instead, the structure may be controlled in cluster units.

Can nanostructure science and technology be promising subjects for chalcogenide glasses? Until now, several scientific problems have attracted considerable interest, while applications have been limited. Accordingly, the answer remains. We can see, however, that, in comparison with oxide and halide glasses, chalcogenide glass has two characteristic features, that is, low-dimensional (molecular) structures and semiconductor properties [1, 2], which provide unique subjects for nanostructure science and technology as follows.

The low-dimensional structure presents several topics. First, the chalcogenide glass shows clear signatures of the medium-range structural order of 1–5 nm, while real structures are still controversial [1, 2]. Accordingly, artificially produced chalcogenide nanostructures could be investigated for obtaining insight into the medium-range structure. For instance, it will be interesting to examine how the first sharp diffraction peak and the boson peak [1, 2] change with decreasing size of the chalcogenide glasses. Second, as shown in Figure 1, the molecular structure has nanometer scales, so that the chalcogenide could be a good test material for investigating the relationship among molecules, crystals, and bulk glasses. How the glass transition and thermal crystallization behave in nanoscale chalcogenides is an intriguing problem. Third, the low-dimensional structure can have a few dangling bonds at the surface, which may be favorable for examining some inherent properties in nanostructures.

As a semiconductor having an energy gap of 1–3 eV, the chalcogenide glass may show notable electronic properties when nanostructured. One is the quantum effect, which is familiar in nanostructured crystalline semiconductors. Also, in a-Si:H (hydrogenated amorphous Si), some quantum effects have been reported, and, accordingly, comparative studies for chalcogenide glasses are important. The other is the photoinduced phenomenon. It is known that chalcogenide glasses exhibit a variety of photoinduced phenomena [3–5], and the behavior in nanostructured samples is interesting from fundamental and applied points of views.

This review is intended to be a summary of nanostructured chalcogenide glasses at the present stage and could also be an embryo for future studies. Section 2 focuses on nanoscopic inspections of chalcogenide glasses. Sections 3–5

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (629–640)

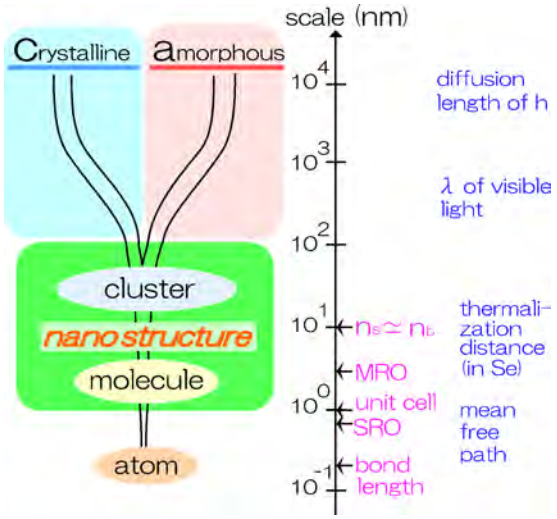


Figure 1. Schematic definition of a nanostructure. Related lengths are also shown. SRO, short-range order; MRO, medium-range order; n_s, number of atoms at the surface: n_b, number of atoms in the interior.

are organized according to the form of the nanostructures: Section 3, zero- and one-dimensional molecular structures; Section 4, two-dimensional nanostructures such as ultrathin films and multilayers; and Section 5, nanostructures produced by modification with radiation and by using scanning-probe microscopy. Since phase-change memories, the so-called Ovonic memories, are growing in use as DVDs (digital video/versatile discs) and the mark size is decreasing to the nanoscale, this topic is treated separately in Section 6. The last section gives a future perspective.

2. NANOSCOPIC INSPECTION

2.1. Transmission Electron Microscopy and Scanning Electron Microscopy

Direct inspections of nanometer-scale structures in chalcogenide films have been carried out using TEM (transmission electron microscopy) and SEM (scanning electron microscopy). Several studies have taken images of inhomogeneous structures greater than about 10 nm in evaporated As_2S_3 and $GeS(Se)_2$ films [6–10]. For instance, in obliquely deposited films, slanted columnar structures have been imaged [11–13], as shown in Figure 2 [10]. The slanted angle changes with the deposition angle, following the so-called tangent law, $\tan\alpha = 2\tan\beta$ [14], where α and β are the vapor incidence angle and the column inclination angle, both measured with respect to the substrate normal. The oblique deposited films may have anisotropic physical properties such as optical birefringence, which remain to be studied.

However, we should be careful of electron-beam effects in TEM and SEM observations. For instance, the small electrical conductivity of chalcogenide glasses may cause image distortions due to charge accumulation. In addition, the chalcogenide glass is liable to undergo structural changes upon irradiation of electron beams [15], and, accordingly, fresh images may not be obtained.

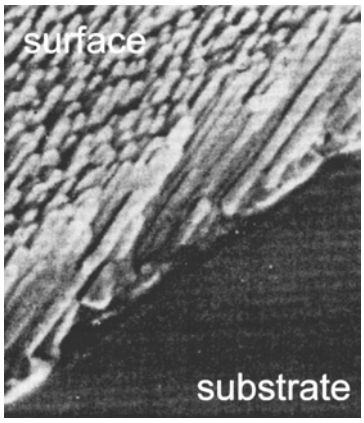


Figure 2. SEM image of an obliquely deposited 1-μm-thick As_2S_3 film [10], which is prepared with a vapor incidence angle of 80°. Note that the column tilt angle is about 45°. Adapted from [10], N. Starbov et at., *J. Non-Cryst. Solids* 139, 222 (1992).

2.2. Scanning Tunneling Microscopy and Atomic Force Microscopy

The surface structures of amorphous chalcogenides have been investigated using STM (scanning tunneling microscopy) and AFM (atomic force microscopy) with nanometer resolution. Note that, in contrast to oxide glasses, some chalcogenide glasses have finite electrical conductivity ($\geq 10^{-6}$ S/cm), which is needed for STM inspections. Ichikawa has examined $Ge_{20}Te_{80}$ bulk glass using STM [16] and AFM [17] at 1 atm. The glass surface is disordered, as expected, while an annealed sample at 100 °C (T_g = 160 °C) for 1–3 days shows atomically ordered regions and some structures that may correspond to a medium-range order. Peled et al. investigated a-Se and its crystallization behavior under laser heating using AFM at 1 atm [18]. They showed an image that may be identified with disordered chain molecules with a scale of 1 nm. Tanaka and co-workers studied the initial stages of photosurface deposition [4] in $Ag_{20}As_{32}Se_{48}$ films, which were exposed to 633 nm light, using STM [19]. Interestingly, in contrast to previous macroscopic observations [4], Ag nanocrystals grow upon illumination in such an Ag-deficient glass. Scanning tunneling spectroscopy with nanoscale spatial resolution has also demonstrated nanoscopic movements of Ag ions in Ag–As–Se glasses [20, 21].

In these studies, however, it may be fair to state that *atomic* images of glass surfaces have not been obtained. The reason is ascribed to the disordered structures. As illustrated in Figure 3, it is plausible that the surface of amorphous materials is irregular with a typical roughness of 1–5 nm, the scale of medium-range orders, and the apex of the STM tips has a curvature of about 5 nm. In such cases, atomic images cannot be obtained, since the tunneling current does not necessarily flow from an apex atom of the tip. Clustering of electron wave functions in disordered glass surfaces may also limit atomic resolution [22]. To obtain atomic images, we need atomically flat *facets*, the size being much greater than the dimension of the tip apex. However, the existence of such facets must be very exceptional in disordered structures [23]. In addition, it is difficult to distinguish real amorphous

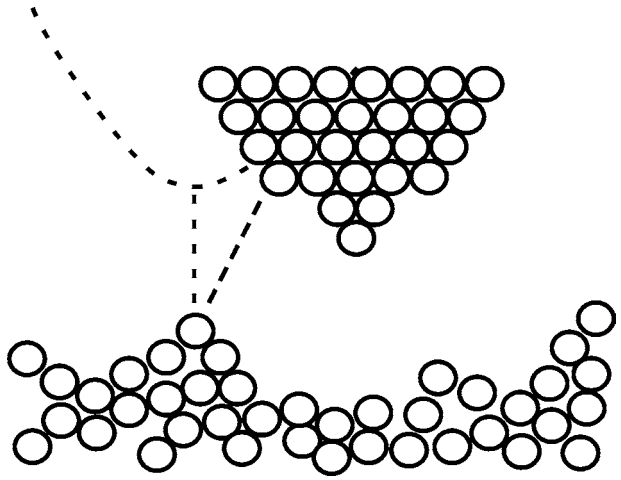


Figure 3. Schematic illustration of a crystalline STM tip and a clean surface of amorphous materials.

structures from contaminated structures, since we cannot identify atomic species using STM. Both real amorphous structures and contaminated surfaces may appear as irregular. Therefore, it will be difficult to get convincing insight into the atomic structure of chalcogenide glass surfaces.

However, it may be valuable to note that, in some molecular and crystalline chalcogenides, clear atomic images have been obtained at 1 atm. Peled et al. presented an AFM image of trigonal Se [18], and Ohto and Tanaka provided STM images of crystalline Ag_2Te surfaces in ordered (20 °C) and superionic (170 °C) phases [24]. Czajka et al. studied the structure of vacuum-evaporated Se molecules onto highly oriented pyrolytic graphite (HOPG) substrates (see the next section) [25].

3. ZERO- AND ONE-DIMENSIONAL MOLECULAR STRUCTURES

Dot structures embedded in wide-gap materials have been studied for chalcogenide materials. However, for amorphous chalcogenide dots, little work has been reported. For instance, Yumashev et al. studied induced optical absorption in a-Se particles (~20 nm) embedded in a polymer [26]. On the other hand, substantial studies have been reported for crystalline chalcogenide dots, for example, CdS(Se) particles in SiO_2 matrices (see, e.g., [27–30]). These studies concern mostly optical properties, which may be promising for filters and nonlinear devices. CdS(Se) crystals have direct bandgaps, in which the optical absorption edges are sharp and excitons can easily be excited. However, the atomic bonding structure is tetrahedral, and, accordingly, when the crystals are nanostructured, many dangling bonds are liable to exist at the surface. On the other hand, in materials such as Se and As_2S_3, in which chalcogen atoms are twofold coordinated, such surface defects can be fewer, so that the intrinsic dot properties can be investigated.

Chalcogen molecules incorporated into nanoscale channels and cavities (see Fig. 4) have attracted substantial scientific interest. Physical studies were pioneered by Bogomolov and co-workers at the Ioffe Institute around 1980 [31–35]. Incorporated molecules are Se [33–55], which has

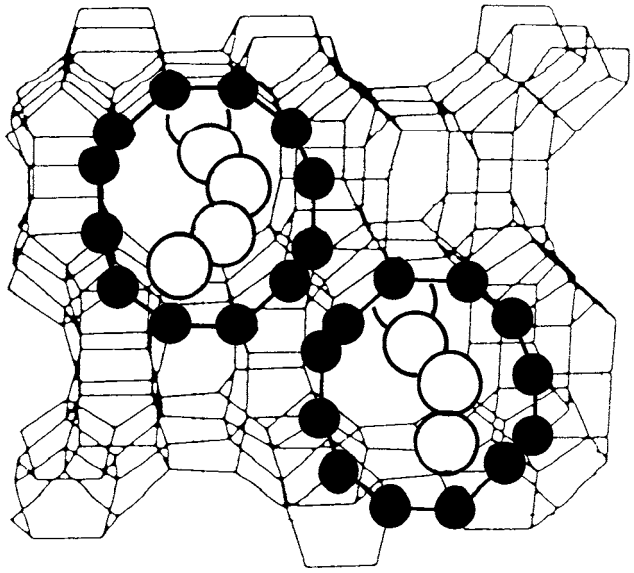


Figure 4. Illustration of Se molecules incorporated in zeolite channels.

been extensively studied, S [31, 50, 54], Te [32, 33, 50], and As–S [56]. On the other hand, porous matrices are mostly a variety of zeolites (crystalline $M_{2/n}O/Al_2O_3/xSiO_2/yH_2O$, where M is a cation such as Na^+, n is the valence number, $x \geq 2$, and $y \geq 0$) [57], and porous glasses [38]. These materials can be prepared in powder form with a size of 0.1–1 μm or as small crystals with a size of 5–50 μm, in which channels and/or cavities with diameters of approximately 1 nm are arranged periodically (zeolites) or nonperiodically (porous glasses). Incorporation of molecules has been done as follows: First, porous materials are dehydrated by heating. Then, the materials are sealed with chalcogenide pellets in an evacuated glass ampoule. Finally, the ampule is heated, which incorporates chalcogenide molecules into the porous structures through thermal diffusion.

Since a variety of zeolites are available (at least 135 kinds) [57], the molecular systems produced are also fond of variety. Structural studies have been done using TEM [37, 43], extended X-ray absorption fine structure (EXAFS) [36, 39, 51, 55], X-ray diffraction [45, 48], Raman scattering [34, 35, 43, 44, 47, 52, 54–56], and electron spin resonance (ESR) [39]. For instance, for Se, not only Se chain molecules, which form the trigonal crystal, but also Se_2 anions [45, 51] and rings such as Se_6, Se_8, and Se_{12} [50, 54] are identified as being incorporated. Photoinduced structural changes have been investigated [34, 39, 42], and some theoretical calculations have also been reported [58–60].

The thermal properties of these molecules are interesting. Matsuishi et al. investigated glass-transition and crystallization behaviors, the results being shown in Figure 5 [49], which will be discussed in Section 4.2.2. Poborocci et al. demonstrated that the molecular shape of Se in a zeolite undergoes a phase transition at certain temperatures [52].

Among many physical properties, optical absorption spectra are the ones most extensively studied [33, 34, 36, 39, 40, 44, 46]. It is amply demonstrated that the optical absorption edge of molecules blueshifts from that of the corresponding amorphous films (see Fig. 6b). This is reasonable since the intermolecular interaction, which governs the width of the

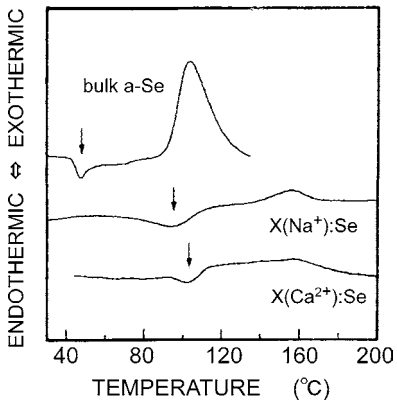


Figure 5. Thermograms obtained using differential scanning calorimetry for amorphous Se (upper) and Se in two kinds of zeolites (lower two) [49]. The arrows indicate the glass transition temperatures.

valence band in chalcogenide glasses [1, 2], is absent in the molecular structures consisting of single molecules. On the other hand, when incorporated into anisotropic matrices such as single crystals of natural mordenite, a kind of zeolite, the optical spectra appear to be highly anisotropic [33]. Studies on photoluminescence [53] and electrical conductivity are relatively limited [32].

The properties of single molecules may be investigated in this way, while the interpretation of experimental results should be done carefully for two reasons. One is a surface effect. That is, in the case of powdered zeolites, chalcogen atoms may be deposited as films also on the powder surface, which will be undesired components, giving rise to noisy properties. The other is a channel effect. For instance, there is direct evidence obtained through TEM images that single Se molecules are incorporated into the narrow channels in zeolite [37]. However, there is also substantial evidence that the molecules interact with the inner walls of the channels [48, 51]. As suggested by theoretical studies, the wall–molecule interaction has a crucial effect on the structural and physical properties [58–60]. Accordingly, if we really want to examine the properties of single molecules, the molecule should be laid on an inert surface or should be produced in free space.

Such studies have been carried out. Czajka et al. prepared, using vacuum evaporation, periodically located monolayers of Se_n rings (where n may be 5–8), the average diameter being approximately 0.5 nm, onto HOPG substrates [25]. STM inspection at 1 atm demonstrated that the ring plane is parallel to the HOPG substrates. The molecules were stable when imaged under tip voltages of −0.7–+0.3 V. However, application of +0.7 V to the STM tip collapsed the molecules, producing oval and round-shaped islands with a diameter of about 5 nm. They interpreted this phenomenon as electronically induced molecular breaking [61], while no explanations were given for, for example, the polarity dependence. On the other hand, Nagata et al. produced neutral clusters of Se_2, Se_5, and Se_6 in vacuum using a supersonic jet expansion method [62]. They found that Se_5 is unstable when ionized by electrons of 14 eV and irradiated by laser light with an intensity of 0.7 W and a wavelength of 488 nm. The instability of odd-membered rings seems to be consistent with theoretical calculations [61, 62]. It is mentioned

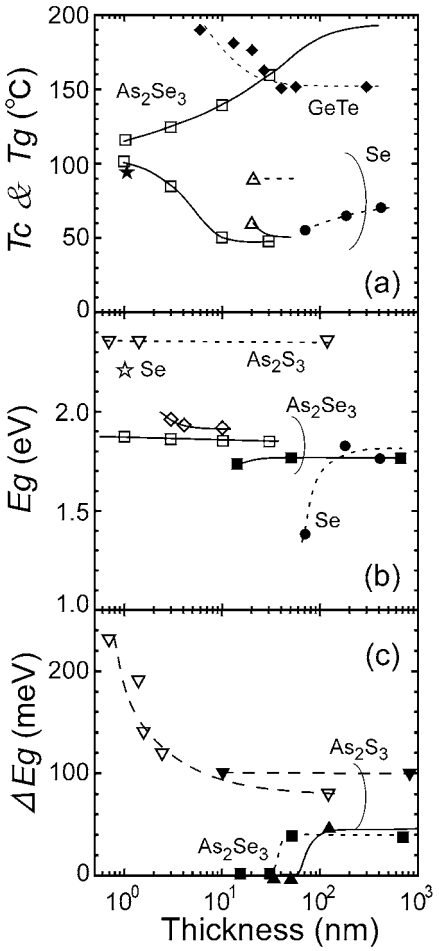


Figure 6. Dependence of thermal and optical properties on the sample dimension such as film thickness and multilayer period. (a) shows the crystallization temperature T_c (dashed lines) of Se and GeTe and the glass transition temperature T_g (solid lines) of Se and As_2S_3. (b) shows the optical gap E_g of Se, As_2Se_3, and As_2Se_3. (c) shows the reversible photodarkening ΔE_g in As_2S_3 at room temperature (solid line) and 80 K (dashed lines) and in As_2Se_3 at room temperature (dotted line). Solid symbols are obtained for films (● [64], ■ [72], ▲ [68], ▼ [71]), open symbols are for multilayers (▽ [75], □ [79], ◇ [80], Δ [115]), and stars (☆ [36], ★ [49]) are for Se in zeolites. Since the details of measurements are different among the results cited, absolute magnitudes are less valuable, and only the trends should be compared.

here that nanoscale powders of a- and c-Se have been prepared through evaporation of Se in Ar gas [63]. The work is intended to show the formation mechanism of amorphous interstellar grains.

4. TWO-DIMENSIONAL NANOSTRUCTURES

4.1. Ultrathin Films

Several kinds of ultrathin films have been prepared, and the thickness dependence of the physical properties has been investigated. These include crystallization kinetics [64, 65], electrical resistance [66], optical absorption [64, 67–72], and photodarkening [69–72]. For instance, Chaudhuri et al. [64]

showed for a-Se films with thickness of 70–180 nm that thermal crystallization occurs at lower temperatures in thinner films (see Fig. 6a). In addition, as shown in Figure 6b, the optical gap decreases from 1.8 to 1.4 eV with a decrease in the film thickness. They interpret these observations as resulting from higher defect densities in thinner films. On the other hand, Chopra and Bahl demonstrated that the crystallization temperature in GeTe films increases in thinner films [65]. These characteristics are discussed in comparison with those in multilayers in Section 4.2. For ultrathin films, however, we should be careful of surface effects, such as oxidation and moisture [73]. Also, it is plausible that the substrates have appreciable effects [10, 69, 74].

The dependence of the photodarkening on film thickness is intriguing, as summarized in Figure 6c. Tanaka and co-workers found that reversible photodarkening [3] disappears in As_2S_3 films thinner than approximately 50 nm when illuminated at room temperature by bandgap light [68]. At about 80 K, it appears as demonstrated in successive studies [70–72]. Nonstoichiometric films ($As_{43}S_{57}$ and $As_{20}S_{80}$) do not show such anomalous thickness dependence [68]. The room-temperature characteristic has been confirmed recently by Hayashi and Mitsuishi using As_2Se_3 films with a thickness of 15–700 nm [72]. They also found that the thickness dependence arises from the initial state, that is, the annealed state before illumination. That is, the optical absorption edge of annealed films redshifts in thinner films (see Fig. 6b), which may be consistent with the result reported by Chaudhuri et al. [64]. Hayashi and Mitsuishi proposed a model as follows [72]: Since the films are attached on oxide glass substrates and the film/substrate combination is annealed before illumination, thermal stresses, which can darken the film, arise. Such a mechanism is possibly more effective in thinner films, so that the photodarkening becomes smaller. On the contrary, Indutnyi et al. [75, 76] showed that much thinner As_2S_3 films (0.7–2.5 nm in thickness), which are embedded in transparent SiO layers, give greater (∼3 times) photodarkening at 80 K than that in thick films. In their result, the annealed state is not dependent on the film thickness (see Fig. 6b), and, instead, the illuminated state changes with the thickness. They interpret this result as arising from the higher density of self-trapped excitons in thinner films, which is assumed to enhance the photodarkening. It is also mentioned that Sainov et al. have reported that refractive-index changes can be photoinduced in thin (10–20 nm) As_2S_3 films [77]. Further studies remain.

4.2. Multilayers

4.2.1. Historical

After the pioneering work on semiconductor superlattices by Esaki and Tsu [78], similar structures have been explored for amorphous semiconductors. Since the constraint of unit-cell matching in crystalline superlattices does not exist in amorphous structures, preparations of amorphous multilayers by repeating depositions are straightforward. However, the boundary between two layers in amorphous multilayers cannot be atomically flat, due to disordered structures, which may cause some unique features. Fundamental problems concern the formation and stability of multilayer structures upon thermal treatment and light illumination. Quantum effects also become to be interesting topics.

As far as we know the first study on amorphous-semiconductor multilayers was reported in the beginning of the 1980s for chalcogenide glasses in Japan, independently from two groups at Hitachi [79] and NTT (Nippon Telephone and Telegram Corp.) [80]. The Hitachi group prepared Se/As_2Se_3 multilayers through vacuum evaporation in the development of highly sensitive vidicon targets. The NTT group tried to see, using sputtered $As_2Se_3/GeSe_3$ multilayers, what occurs when long-range multilayer periodicity is incorporated into amorphous semiconductors. They measured several physical properties such as optical absorption, electrical conductivity, glass-transition temperature, and infrared (IR) spectra.

Afterwards, the work by Abeles and Tiedje [81] triggered extensive studies on a-Si:H, which may stimulate further investigations of chalcogenide multilayers. Trodal and Wright prepared Ge/Se multilayers, and discovered photoinduced diffusion of Se atoms into Ge using Raman scattering spectroscopy [82, 83].

However, in these studies of chalcogenide multilayers, it was not explicitly demonstrated if long-range structural periodicity was really formed. Since the initial stage of evaporation starts with the formation of island structures due to the surface tension of the deposited clusters [74, 83], repeated deposition cannot necessarily produce multilayers. In addition, there is a possibility that interdiffusion occurs at multilayer interfaces, so that homogeneous structures may be produced [85]. Actually, Tanaka et al. tried to prepare multilayers of Se/As_2S_3, but they could not detect X-ray Bragg peaks originating from multilayer periodicity [86]. Proper selection of materials and deposition conditions is required.

The first structural evidence of chalcogenide multilayers was given by Shirai et al. for Se/Cd–Se multilayers with periodicities of 4–40 nm [87]. Later, a Bulgarian group obtained similar results for evaporated Se/CdSe multilayers [88], which are shown in Figure 7. In these studies, the researchers detected clear X-ray Bragg peaks originating from the long-range periodicity, when the layers are thicker than 2–3 nm. These results mean that the thickness of interfacial regions can be estimated at 1–2 nm. We note here that in ideal crystalline superlattices, the interfacial thickness is nominally equal to the atomic distance, approximately 0.2 nm. On the other hand, in amorphous multilayers, we can assume that, due to disordered structures, rough interfaces and interfacial strains necessarily produce the interfacial region of 1–2 nm thickness.

The Bulgarian group has continued the most active research on chalcogenide multilayers using Se(Te)/CdSe and SeTe/Se combinations [89–104]. In addition to thermal vacuum evaporation systems having rotating substrate holders, they have also employed laser ablation [100].

Others have studied some properties of several kinds of multilayers. For instance, Imamura et al. investigated carrier transport in SeTe/Se multilayers [105], and Hamanaka et al. measured photoluminescence in As_2Se_3/As_2S_8 [106]. For photoinduced interdiffusion, Se/As_2S_3 [107–111], $SeTe/As_2S_3$ [112], and Ag/AsS_2 [113] have been examined. Here, the last combination has shown a substantial increase in photodoping efficiency [4] by multilayering, which may

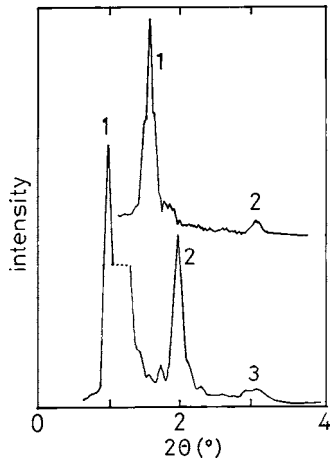


Figure 7. X-ray Bragg peaks obtained for Se/CdSe multilayers with periodicity of 5.7 nm (upper) and 9.2 nm (lower) [89]. The number indicates the diffraction index. In the lower trace, the fundamental peak is weakened.

be promising for some applications. Indutnyi et al. [75, 76] investigated the photodarkening in As_2S_3/SiO as already noted in Section 4.1, and Kim et al. measured thermal boundary resistance in $Ge_2Sb_2Te_5$/ZnS multilayers with an emphasis on to phase-change memories [114].

4.2.2. Thermal Stability

Studies on thermal stability are limited, probably because the sensitivity of thermal analyses is not sufficient for thin films. Actually, a few direct thermal analyses have been reported [49, 79], while indirect analyses using structural and optical measurements also provide additional insight [103, 115]. Vateva investigated thermal modifications of Se(Te)/CdSe multilayer structures using X-ray diffraction [100]. Thermal annealing can provide conventional intermixing (positive interdiffusion) as well as phase separation (negative interdiffusion) at multilayer interfaces, which occur depending on several conditions such as annealing temperatures and periods [100].

Results for the thermal crystallization are summarized in Figure 6a. We see that the observations appear to be controversial. Popescu et al. showed that, in the multilayered samples of Se/CdSe with a layer periodicity of about 20 nm, the crystallization temperature (90–100 °C) of Se is similar to that in bulk Se [115]. In contrast, it can be pointed out that no clear Se crystallization is detected in Se/As_2Se_3 multilayers with periodicities of 1–30 nm [79] and in Se/zeolite (see Fig. 5) [49]. For the crystallization to occur, the scale of the Se films (clusters) should be greater than the size of the crystallites or should be much greater than the crystalline unit-cell dimension, which is approximately 1 nm. These conditions may not be satisfied in these two kinds of nanostructured samples, and, accordingly, Se crystallization cannot occur. On the contrary, it has already been mentioned in Section 4.1 that, in Se films with a thickness of 70–180 nm, the crystallization occurs at lower temperatures [64]. Such thermal instability is commonly observed for the melting temperature T_m, which is understood theoretically using the Thomson–Frenkel relation given as $T_m - T_m(t) \propto 1/t$, where t is the thickness. The weaker bonding of atoms located at and near the free surface is assumed to govern these decreases in characteristic temperatures. How can we reconcile these contrasting behaviors? It is plausible that the difference between multilayers and films, that is, the existence of interfacial layers, has some effect on these characteristics [103].

For the glass transition, as shown in Figure 6a, the results also appear to be diverse. In the Se/As_2Se_3 multilayers [79] and Se/zeolite [49] samples, the glass-transition temperature of Se increases in more spatially restricted samples. On the other hand, the temperature of As_2Se_3 in the Se/As_2Se_3 decreases in samples with smaller periodicities. In addition, it is interesting to note that, in the multilayer samples, the temperature increase in Se and the decrease in As_2Se_3 occur at a thickness of approximately 50 nm [79].

The glass transition is still a matter of controversy in solid-state science, and this behavior remains to be studied. Recent studies on other materials in confined geometries show that T_g increases in more confined samples if there are no interfacial interactions for the confinement. This behavior is understood as a manifestation of collective molecular motions of the glass transition [116–118].

4.2.3. Electronic Quantum Effect

Although multilayer structures have been demonstrated in X-ray diffraction patterns [87, 88, 94], phonon spectra [92], and SEM images [106], electronic quantum-size effects still remain vague. It is well known that an electron confined in a one-dimensional square well with infinite potential walls has an energy of $E_n = (\hbar^2/2m)(n\pi/L)^2$, where m is the electron mass, n the positive integer, and L the well thickness. Accordingly, we may expect blueshifts ($\propto 1/L^2$) of the optical absorption edges and stepwise ($\propto n^2$) absorption spectra in thin films. Actually, as shown in Figure 6b, in some systems the blueshifts of the optical absorption edges are reported [80–82, 87]. In others, however, the edges do not shift [75, 79]. In addition, we have already seen in Section 3 that ultrathin Se and As_2Se_3 films even exhibit red shifts when the film thickness is decreased [64, 72]. For the stepwise optical spectra, the Bulgarian group reported some traces in SeTe/CdS and SeTe/Se systems having a periodicity of 4–7 nm [92, 94, 100].

However, when interpreting these optical results, we should be aware of several points. First, since we are dealing with thin films, optical interference fringes inevitably appear on optical transmission spectra. This optical effect should be clearly separated when evaluating absorption spectra. Second, for the blueshift, interfacial layers (~2 nm thick) can be responsible, which is also inferred from carrier-transport behavior [87, 105]. This structural effect becomes progressively greater in thinner films, and, accordingly, it is very difficult to separate the quantum effect from the interfacial effects. Third, we should note that, in formulations of the quantum effect, it is implicitly assumed that the mean free paths of electrons and/or holes are much longer than the well thickness. Taking the small mobility in amorphous semiconductors into account, we can assume that the mean free path is, at most, 2 nm [119], which is comparable to the interfacial layer thickness [88, 94]. Therefore, it seems difficult to obtain unambiguous signatures of the quantum effects. Fourth, for the case of SeTe/CdS multilayers [93],

it is a pnpn-type system. That is, electrons in the valence band in SeTe should be excited into the conduction band in CdS. The two states are spatially separated, and, accordingly, the transition probability must be very small. Such a situation probably makes the observation of quantum effects nearly impossible. As far as the author knows, among many studies of amorphous semiconductor multilayers, including chalcogenide and tetrahedral materials, the clearest quantized (stepwise) absorption spectrum is the one obtained by Hattori et al. [120] for a-Si:H/a-SiC:H with layer thicknesses of 2–5 nm at 100 K.

5. STRUCTURAL MODIFICATIONS

5.1. Radiation Effects

Chalcogenide nanostructures can be produced using a variety of radiation such as X-ray, electron beams, and ion beams. Soft X-ray exposures upon Ge–Se films have produced microscopic structures with scales down to 10 nm [121–123] through photodoping processes [4]. On the other hand, electron [124–127] and ion beams [127] can produce nanostructures in chalcogenide films such as Ag/As–S without using patterned masks. These studies aim at developing novel lithographic techniques having ultrahigh resolutions.

For visible light, nanostructures can be prepared using near-field optics. Such trials have been reported for photodarkening [3], photochemical modification [4], and phase-change phenomena, the last being treated separately in Section 6.1. For photodarkening, Isbi et al. demonstrated, using $As_{50}Se_{50}$ films and an SNOM (scanning near-field optical microscope), that dark lines of about 100 nm width can be recorded using green light [128]. They also mentioned that Al-coated fiber tips are liable to break upon light illumination of 10 nW, which corresponds to 1 W/mm^2, through thermal vaporization of the Al film. For photochemical modification, Gotoh and Tanaka showed, using Ag–As(Ge)–S and an SNOM providing 0.1–1 nW of blue-laser illumination, that convex deformations with a diameter of 100–300 nm can be produced [129, 130]. The deformation is assumed to be a manifestation of the accumulation of Ag^+ ions [4].

5.2. STM-Produced Structures

STM is capable of producing surface modifications with nanoscale resolution at *room temperature* [131]. The modification may be topological or compositional, in which case continuous compositional modifications can be induced, in principle, only in amorphous materials such as chalcogenide glasses, which are thermodynamically quasistable [1, 2]. Since some chalcogenide glasses containing Se and Te possess proper electronic conductivity, and Ag-chalcogenide glasses have finite ionic conduction, the nanoscale STM modification appears to be an interesting subject. (Note that these features are in contrast to atomic manipulation of crystalline surfaces, which can be performed in vacuum at low temperatures.) However, the application of STM modifications of chalcogenide glasses has been rather limited. Comparative studies using electrical AFM may also be valuable. The work on Se molecules [25] has already been described in Section 3.

The first demonstration of nanoscale modifications of chalcogenide films was reported by Utsugi [132]. He showed, using $GeSe_9$ films coated with Ag–Se, that STM is capable of producing grooves approximately 10 nm wide and 3 nm deep through an electrochemical reaction. That is, when the bilayer film consisting of Ag–Se/$GeSe_9$ is subjected to tip voltages of about +4 V, Ag^+ ions in Ag–Se layers migrate into $GeSe_9$, and, as a result, Se atoms remain at the surface. Then, if the sample is treated in H_2 gas ambient, the Se atoms react to form H_2Se, which evaporates, giving rise to the narrow grooves. Using this phenomenon, he could draw nanoscale patterns such as shown in Figure 8. Later, Ohto and Tanaka reported that, also in Ag–As–Se glass, Ag^+ ions migrate under STM electric fields [133], which is consistent with the observation by Utsugi.

Different processes of STM modifications appear in other materials. In Cu–As–Se glasses, as shown in Figure 9, sombrero-shaped modifications have been produced by applying negative tip voltages of about −5 V [133, 134]. In contrast, Cu_3AsSe_4 crystal does not show such deformations, the phenomenon being inherent to glasses. It seems that injected electrons from the tip to the glass induce appreciable fluidity, which may be similar to the fluidity induced by electron beams [15] and light [135]. In addition, Coulomb attractive forces between the tip and the glass possibly govern the shape, which resembles that of a Taylor cone. Curiously, in As_2Te_3, which has a similar electrical conductivity to that of Cu–As–Se, only conventional field evaporation was induced under tip voltages higher than ±6 V [134]. The reasons of these different features remain to be studied.

6. OVONIC MEMORIES

Optical phase-change memories were commercialized in the middle of the 1990s, which has accelerated intensive research and development. The memory utilizes reversible structural changes between amorphous and crystalline phases in telluride films such as Ge–Sb–Te, which are

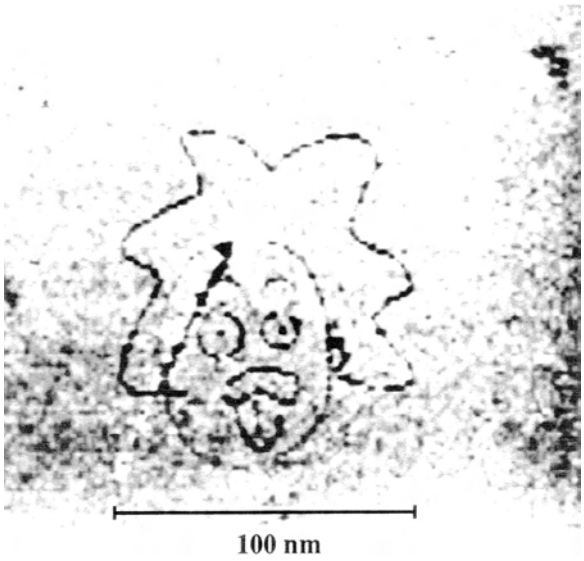


Figure 8. Smallest Einstein. Courtesy of Y. Utsugi.

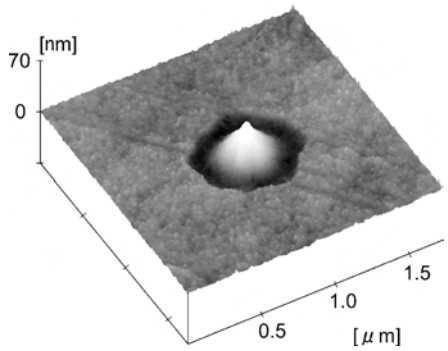


Figure 9. A sombrero structure in $Cu_{30}As_{28}Se_{42}$ glass after an application of −5 V for 1 s with an STM tip [134].

induced by temperature rises produced by light pulses. The bit size is now submicrometer, which is decreasing to the nanoscale [136]. Film structures were inspected using TEM with 1–10 nm resolution [137] and AFM [137, 138].

These studies shed light, not only on applications, but also on fundamental problems. For instance, what are the smallest amorphous or crystalline marks that can be written in phase-change films? We can assume that, if the dimensions are the same, crystalline spots in amorphous phases are more stable than amorphous spots in crystalline phases, since the crystal has a smaller free energy. Therefore, in a feature, alternatively to the present scheme, the crystalline mark may be written and erased in amorphous films.

6.1. Optical

At least two trials have been reported for nanoscale optical phase-change recordings. Hosaka et al. demonstrated the production of 60-nm marks in 30-nm-thick amorphous $Ge_2Sb_2Te_5$ films using an SNOM and 785-nm light pulses with widths of 0.5 and 5 ms [139]. Written marks are read as reflection SNOM images. They expect that a high-speed reading of 10–100 Mbit/s will be achieved. On the other hand, Tominaga and co-workers produced 90-nm marks in $Ge_2Sb_2Te_5$ films using 686-nm light in combination with a nonlinear transmission phenomenon in thin Sb films [140, 141]. Their sample consists of a triple-layer structure, Sb(15 nm)/SiN(20 nm)/$Ge_2Sb_2Te_5$(15 nm). Then, intense laser pulses produce optical pinholes in the Sb film with a diameter smaller than the light wavelength, and evanescent light from the film induces a phase change in the chalcogenide film. In this experiment, the Sb film works as a nonlinear optical switch, while its behavior has not been analyzed.

Here, we should make two points. One is that, in these studies, as-deposited amorphous films are employed for recording, and amorphous-to-crystalline phase changes are induced. The other is that mark erasing has not yet been demonstrated, which is probably more difficult, since we should manage more intense light pulses. However, intense light may destroy SNOM tips or produce larger optical pinholes. Therefore, much work will be needed for practical applications of these optical nanoscale memories.

6.2. Electrical

Electrical Ovonic devices of nonvolatile RAM (random access memory) type are currently being revived. The Ovshinsky group reported on a new memory utilizing Ge–Sb–Te films 50 nm thick, in which active bit sizes are as small as 200 nm [142, 143]. Nakayama et al. employed As–Sb–Te films 200 nm thick, in which the bit size is 300 nm [144]. These bit sizes are determined predominantly by bit-cell lithography techniques, and further reduction will be possible.

On the other hand, electrical Ovonic memories utilizing AFM have also been explored. Kado and Tohda tried to induce phase changes using an AFM having conducting cantilevers [145]. Their samples have a two-layer structure consisting of $GeSb_2Te_4$ (20 nm thick) and Pt films (200 nm thick), which are sputtered onto Si wafers. Mark writing can be made by applying positive pulses with widths of 10 μs–5 ms, voltages of 2–10 V, and currents of 10–20 nA to the $GeSb_2Te_4$ film through Au-coated Si_3N_4 cantilevers. Written data are read under a positive direct-current (dc) bias of 0.5–1 V. Drastic changes (∼100 times) in electrical conductivity of the film have been detected, which accompany no topological changes, within spots as small as 10 nm. This size corresponds to a data storage density of 1 Tbit/cm^2. In a more recent work [146], they succeeded in mark erasing by scanning the cantilever, which applies −1 V. However, this polarity dependence, that is, writing by positive voltages and erasing by negative voltages, strongly suggests that the change is electronic, not thermal, accompanying structural phase changes [146]. Actually, Gotoh et al. demonstrated recently that such marks are unstable, spontaneously disappearing within 1 day [147].

Gotoh et al. also demonstrated that erasable amorphous-to-crystalline phase changes are possible using similar samples to that of Kado and Tohda [147]. An important difference is that their cantilevers are made from PtIr wires, which permit high-current flows. Then, crystalline marks with a diameter of approximately 50 nm can be written by electrical pulses of 1.5 V (∼1 mA) and 2 μs and erased by 2 V and 20 ns (see Fig. 10). No polarity dependence has been observed for the writing and erasing. In addition,

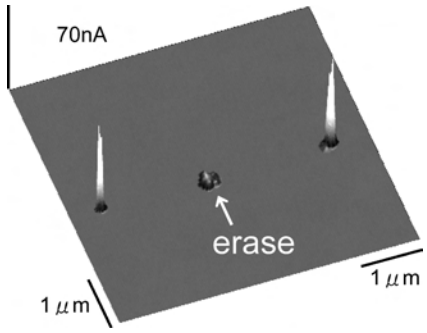


Figure 10. Erasable nanoscale phase changes produced by electrical AFM in a 50-nm-thick $GeSb_2Te_4$ film. In the three marks, the left hand and right hand are written by single pulses of 1.5 V and 2 μs, while the central mark is written by the pulse and then erased by a pulse of 2 V and 20 ns. This current image is obtained by luster scanning a conductive cantilever with 10 mV. Courtesy of T. Gotoh.

X-ray diffraction measurements detected crystalline peaks for written marks. These facts show that structural phase changes governed by Joule heats are electrically induced. This may be the first demonstration of *erasable nanoscale* Ovonic memories.

Similar trials have been done using STM, although the results are still preliminary. Saluel et al. demonstrated, using 30-nm-thick Ge–Sb–Te fillms, that voltage pulses greater than 3 V and 10 ns produce holes with a diameter of approximately 200 nm [148]. On the other hand, Sugawara et al. applied dc voltages to $GeSb_2Te_4$/Pt/Si samples [149]. Three kinds of deformations are produced. That is, the tip voltage of -10 V gives prominent expansions, $+2$ V gives depressions with peripheral mounds, and $+10$ V gives depressions with no peripheral deformations. The shape of depressions by $+2$ V suggests that these are produced by tip scratching. On the other hand, the ± 10-V deformations appear only when ambient humidity is higher than approximately 50%, which implies some roles of thin ($\sim$1 nm) water layers on the $GeSb_2Te_4$ films. Anodic oxidation and electrochemical decomposition seem to be responsible for the expansion and the depression.

7. CONCLUSIONS

Intensive motivation is needed for advancing nanostructural studies, since the experiments are more or less difficult due to limited sample sizes. Among such motivation,

Table 1. A variety of chalcogenide glass (molecules and clusters) nanostructures and related topics.

Nanostructure form	Synthetic method	Material	Scale (nm)	Related topic/investigated property	Ref.
Molecular	Dispersion	Se	20	Optical	[26]
	Injection into zeolite	S	1	Raman	[31, 50, 54]
		Se	1	Structural, ESR, PI, thermal,	[33–55]
		Te	1	optical, PL, Raman	
				Optical, electrical, Raman	[32, 33, 50]
		As–S	1	Raman	[56]
	Molecular deposition	Se	1	STM, stability	[25, 62]
	Gas evaporation	Se	$\sim 10^2$	Structural	[63]
Ultrathin film	Evaporation	Se	70–420	Thermal, optical	[64]
		GeTe	6–300	Thermal, electrical, optical	[65–67]
		As_2S_3	10–	PI	[68, 70–72]
		As_2Se_3	15–700	PI	[69]
Anisotropic film	Oblique evaporation	As_2S_3	$\sim$10	Structural	[10–12]
		$GeS(Se)_2$	$\sim$10	Structural	[13]
Multilayer	Repeated evaporation	As_2S_3/SiO	1–3	Optical, PI	[75, 76]
		Se/As_2Se_3	1–30	Thermal, optical	[79]
		Se/Cd–Se	4–40	Structural, optical, electrical	[87]
		Se(Te)/CdSe	3–10	Thermal, structural, Raman, electrical	[88–101, 115]
		Se/SeTe	3–10	optical, electrical, Raman	[95, 103, 105]
		Se/As_2S_3	3–10	PI	[86, 107–111]
		As_2Se_3/As_2S_8	75, 225	PL, SEM	[106]
		Ge/Se	1–10	Raman, optical, PI	[82, 83]
		Ag/AsS_2	$\sim$100	PI	[113]
	Repeated sputtering	As_2Se_3/$GeSe_3$	3–	Optical, electrical, IR	[80]
		$Ge_2Sb_2Te_5$/ZnS	$\sim$20	Thermal	[114]
Modification	Radiation	$Ge_2Sb_2Te_5$	90	PC	[140, 141]
		Ag_2Se/Ge–Se	10–50	X-ray lithography	[121–123]
		Ag_2S/As_2S_3	30	Electron-beam lithography	[124–126]
		$GeSe_3$	100	Fon-beam lithography	[127]
	Electrical	Ge–Sb–Te	200	PC	[142, 143]
		As–Sb–Te	300	PC	[144]
	STM	Ag_2Se/Ge–Se	$\sim$10	Ionic lithography	[131, 132]
		As_2Te_3	$\sim$200	Patterning	[134]
		Ag–As–Se	$\sim$200	Patterning	[133, 134]
		Cu–As–Se	$\sim$200	Patterning	[133, 134]
		Ge–Sb–Te	$\sim$100	Electrical, chemical	[148, 149]
	AFM	Ge–Sb–Te	10	PC	[145–147]
		Ag–In–Sb–Te	$\sim$200	Patterning	[148]
	SNOM	$As_{50}Se_{50}$	100	Patterning	[128]
		Ag–As(Ge)–S	100–300	Patterning	[129, 130]
		$Ge_2Sb_2Te_5$	60	PC	[139]

PC, phase change; PL, photoluminescence; PI, photoinduced changes.

undoubtedly, nanoscale phase-change memories are the biggest at present. The study will produce terabyte memories and also provide fundamental insight into the nanoscale phase change and mark stability. On the other hand, chalcogenide photoresists have been demonstrated to possess 10-nm resolution, which will be competitive with that of conventional organic resists. Nanoscale dots and linear channels of chalcogenides appear to give interesting optical properties such as anisotropy and nonlinearity. The system can be regarded as a kind of artificially controlled nanocomposite, which is analogous to dye molecules in organic polymers.

As for fundamental problems, further studies are need for gaining deeper understanding of the scale dependence. Most observations reported so far are preliminary and, in some cases, controversial. For instance, behaviors of the glass transition remain to be studied in more detail. In addition, if electronic quantum effects really exist in chalcogenide nanostructures is ambiguous. Optical absorption and Raman scattering spectra have been investigated, while few studies have been reported for photoluminescence. As for materials, further studies on Se, which is the simplest chalcogenide system, may be valuable.

Table 1 summarizes the current state of the art of a variety of chalcogenide glass nanostructures.

GLOSSARY

a- Abbreviation for amorphous. The amorphous material is defined as a noncrystalline solid. The glass is a kind of amorphous material, which is produced through quenching of the corresponding liquid. Accordingly, the glass exhibits the glass transition.

c- Abbreviation for crystalline.

T_g The glass transition temperature. The temperature that separates a glass from the corresponding supercooled liquid state. This is not a temperature indicating a phase transition, since the glass is not a thermally equilibrium phase.

ACKNOWLEDGMENT

The author thanks K. Sugawara for preparing the illustrations.

REFERENCES

1. S. R. Elliott, in "Materials Science and Technology" (J. Zarzycki, Ed.), Vol. 9, p. 375. VCH, Weinheim, 1991.
2. K. Tanaka, in "Encyclopedia of Materials: Science and Technology" (K. H. J. Buschow, Ed.), p. 1123. Elsevier, Amsterdam, 2001.
3. A. V. Kolobov and K. Tanaka, in "Handbook of Advanced Electronic and Photonic Materials and Devices" (H. S. Nalwa, Ed.), Vol. 5, p. 47. Academic Press, San Diego, 2001.
4. T. Kawaguchi, K. Tanaka and S. R. Elliott, in "Handbook of Advanced Electronic and Photonic Materials and Devices" (H. S. Nalwa, Ed.) Vol. 5, p. 92. Academic Press, San Diego, 2001.
5. K. Tanaka, in "Handbook of Advanced Electronic and Photonic Materials and Devices" (H. S. Nalwa, Ed.), Vol. 5, p. 120. Academic Press, San Diego, 2001.
6. M. Noda and T. Arizumi, *Jpn. J. Appl. Phys.* 15, 173 (1976).
7. C. H. Chen, *J. Non-Cryst. Solids* 44, 391 (1981).
8. C. H. Chen, J. C. Phillips, P. M. Bridenbaugh, and D. A. Aboav, *J. Non-Cryst. Solids* 65, 1 (1984).
9. A. Singh, *J. Appl. Phys.* 58, 309 (1985).
10. N. Starbov, K. Starbova, and J. Dikova, *J. Non-Cryst. Solids* 139, 222 (1992).
11. B. Singh, S. Rajagopalan, P. K. Bhat, D. L. Pandya, and K. L. Chopra, *J. Non-Cryst. Solids* 35–36, 1053 (1980).
12. K. Starbova, N. Starbov, and J. Dikova, *J. Non-Cryst. Solids* 210, 261 (1997).
13. K. Starbova, V. Mankov, J. Dikova, and N. Starbov, *Vacuum* 53, 441 (1999).
14. H. van Kranenburg and C. Lodder, *Mater. Sci. Eng.* R11, 295 (1994).
15. K. Tanaka, *Appl. Phys. Lett.* 70, 261 (1997).
16. K. Ichikawa, *J. Non-Cryst. Solids* 150, 435 (1992).
17. K. Ichikawa, *J. Phys.: Condens. Matter* 7, L135 (1995).
18. A. Peled, V. Baranauskas, C. Rodrigues, D. Art-Weisman, L. Grantman, and A. A. Friesem, *J. Appl. Phys.* 77, 6208 (1995).
19. K. Tanaka, "Proceedings of the 13th International Symposium on Non-oxide Glasses and New Optical Glasses," 2002, p. 65.
20. Y. Utsugi, *Jpn. J. Appl. Phys.* 32, 2969 (1993).
21. M. Ohto and K. Tanaka, *Appl. Phys. Lett.* 71, 3409 (1997).
22. D. E. Bürgler, C. M. Schmidt, D. M. Schaller, F. Meisinger, T. M. Schaub, A. Baratoff, and H.-J. Günthrodt, *Phys. Rev. B* 59, 10895 (1999).
23. J. Jäckle and K. Kawasaki, *J. Phys.: Condens. Matter* 7, 4351 (1995).
24. M. Ohto and K. Tanaka, *J. Vac. Sci. Technol., B* 14, 3452 (1996).
25. R. Czajka, A. Kasuya, N. Horiguchi, and Y. Nishina, *J. Vac. Sci. Technol., B* 12, 1890 (1994).
26. K. V. Yumashev, V. P. Mikhailov, V. S. Gurin, P. V. Prokoshin, M. V. Artem'ev, and S. P. Zhmako, *Quantum Electron.* 24, 828 (1994).
27. T. Moyo, K. Maruyama, and H. Endo, *J. Phys.: Condens. Matter* 4, 5653 (1992).
28. T. Yanagawa, H. Nakano, Y. Ishida, and K. Kubodera, *Appl. Phys. Lett.* 62, 3414 (1993).
29. D. Nesheva and H. Hofmeister, *Solid State Commun.* 114, 511 (2000).
30. H.-J. Eisier, V. C. Sundar, M. G. Bawendi, M. Walsh, H. I. Smith, and V. Klimov, *Appl. Phys. Lett.* 80, 4614 (2002).
31. K. Seff, *J. Phys. Chem.* 76, 2601 (1972).
32. V. N. Bogomolov, A. I. Zadorzhnii, T. M. Pavlova, V. P. Petrampvskii, V. P. Podkhalyuzin, and A. L. Kholkin, *JETP Lett.* 31, 378 (1980).
33. V. N. Bogomolov, S. V. Kholodkevich, S. G. Romanov, and L. S. Agroskin, *Solid State Commun.* 47, 181 (1983).
34. V. N. Bogomolov, V. V. Poborchii, S. V. Kholodkevich, and S. I. Shagin, *JETP Lett.* 38, 532 (1984).
35. V. N. Bogomolov, V. V. Poborchii, and S. V. Kholodkevich, *JETP Lett.* 42, 517 (1985).
36. K. Tamura, S. Hosokawa, H. Endo, S. Yamasaki, and H. Oyanagi, *J. Phys. Soc. Jpn.* 55, 528 (1986).
37. O. Terasaki, K. Yamazaki, J. M. Thomas, T. Ohsuna, D. Watanabe, J. V. Sanders, and J. C. Barry, *Nature* 330, 58 (1987).
38. C. A. Huber and T. E. Huber, *J. Appl. Phys.* 64, 6588 (1988).
39. Y. Katayama, M. Yao, Y. Ajiro, M. Inui, and H. Endo, *J. Phys. Soc. Jpn.* 58, 1811 (1989).
40. Y. Nozue, T. Kodaira, O. Terasaki, K. Yamazaki, T. Goto, D. Watanabe, and J. M. Thomas, *J. Phys.: Condens. Matter* 2, 5209 (1990).
41. H. Endo, K. Maruyama, T. Tsuzuki, and M. Yao, *Jpn. J. Appl. Phys.* 32, 773 (1993).
42. K. Maruyama, T. Tsuzuki, M. Yao, and H. Endo, *Surf. Rev. Lett.* 3, 711 (1996).
43. P. Armand, A. Goldbach, C. Cramer, R. Csencsits, L. E. Iton, D. L. Price, and M.-L. Saboungi, *J. Non-Cryst. Solids* 205–207, 797 (1996).

44. Z. Lin, Z. Wang, W. Chen, L. Lin, G. Li, Z. Liu, H. Han, and Z. Wang, *Solid State Commun.* 100, 841 (1996).
45. P. Armand, M.-L. Saboungi, D. L. Price, L. Itoh, C. Cramer, and M. Grimsditch, *Phys. Rev. Lett.* 79, 2061 (1997).
46. H. D. Sun, Z. K. Tang, W. M. Zhao, and K. L. Wong, *Appl. Phys. Lett.* 71, 2457 (1997).
47. Z. K. Tang, M. M. T. Loy, T. Goto, J. Chen, and R. Xu, *Solid State Commun.* 101, 333 (1997).
48. V. V. Poborchii, M. Sato, and A. V. Shchukarev, *Solid State Commun.* 103, 649 (1997).
49. K. Matsuishi, K. Nogi, H. Ogura, S. Onari, and T. Arai, *J. Non-Cryst. Solids* 227–230, 799 (1998).
50. V. V. Poborchii, *Solid State Commun.* 107, 513 (1998).
51. A. V. Kolobov, H. Oyanagi, V. V. Poborchii, and K. Tanaka, *Phys. Rev. B* 59, 9035 (1999).
52. V. V. Poborchii, A. V. Kolobov, J. Caro, V. V. Zhuravlev, and K. Tanaka, *Phys. Rev. Lett.* 82, 1955 (1999).
53. L. Hei, Z. X. Shen, G. Gu, L. Qin, and S. H. Tang, *Chem. Phys. Lett.* 300, 504 (1999).
54. V. V. Poborchii, *J. Chem. Phys.* 114, 2707 (2001).
55. A. Goldbach, M.-L. Saboungi, L. E. Itoh, and D. L. Price, *J. Chem. Phys.* 115, 11254 (2001).
56. V. Mitsa and I. Fejsa, *J. Mol. Struct., Part 2* 410–411, 263 (1997).
57. http://www.iza-structure.org.
58. A. Ikawa and H. Fukutome, *J. Non-Cryst. Solids* 117–118, 328 (1990).
59. Z. Blum, S. T. Hyde, and B. W. Ninham, *J. Phys. Chem.* 97, 661 (1993).
60. A. A. Demkov and O. F. Sankey, *J. Phys.: Condens. Matter* 13, 10433 (2001).
61. Z. Q. Li, J. Z. Yu, K. Ohno, B. L. Gu, R. Czajka, A. Kasuya, Y. Nishina, and Y. Kawazoe, *Phys. Rev. B* 52, 1524 (1995).
62. K. Nagaya, T. Hayakawa, M. Yao, and H. Endo, *J. Non-Cryst. Solids* 205–207, 807 (1996).
63. C. Kaito and K. Fujita, *Jpn. J. Appl. Phys.* 25, 496 (1986).
64. S. Chaudhuri, S. K. Biswas, A. Choudhury, and K. Goswami, *J. Non-Cryst. Solids* 46, 171 (1981).
65. K. L. Chopra and S. K. Bahl, *J. Appl. Phys.* 40, 4171 (1969).
66. K. L. Chopra and S. K. Bahl, *J. Appl. Phys.* 40, 4940 (1969).
67. K. L. Chopra and S. K. Bahl, *J. Appl. Phys.* 41, 2196 (1970).
68. K. Tanaka, S. Kyoya, and A. Odajima, *Thin Solid Films* 111, 195 (1984).
69. H. Hoshi, Y. Suzuki, and M. Hirai, *J. Non-Cryst. Solids* 95–96, 749 (1987).
70. H. Eguchi, Y. Suzuki, and M. Hirai, *J. Non-Cryst. Solids* 95–96, 757 (1987).
71. H. Eguchi and M. Hirai, *J. Phys. Soc. Jpn.* 59, 4542 (1990).
72. K. Hayashi and N. Mitshuishi, *J. Non-Cryst. Solids* 299–302, 949 (2002).
73. A. Legros, M.-K. Shi, A. Mouton, and A. Seimani, *J. Appl. Phys.* 78, 3048 (1995).
74. D. D. Thornburg, *Thin Solid Films* 37, 215 (1976).
75. I. Z. Indutnyi and P. E. Shepeljavi, *J. Non-Cryst. Solids* 227–230, 700 (1998).
76. I. Z. Indutnyi, P. E. Shepeliavyi, and V. I. Indutnyi, *Semicond. Phys. Quantum Electron Optoelectron.* 2, 59 (1999).
77. S. Sainov, V. Sainov, and J. Dikova, *J. Optoelectron. Adv. Mater.* 3, 399 (2001).
78. L. Esaki and R. Tsu, *IBM J. Res. Develop.* 14, 61 (1970).
79. A. Maruyama, *Jpn. J. Appl. Phys.* 21, 213 (1982).
80. T. Ogino and Y. Mizushima, *Jpn. J. Appl. Phys.* 22, 1647 (1983).
81. B. Abeles and T. Tiedje, *Phys. Rev. Lett.* 51, 2003 (1983).
82. H. J. Trodahl, M. W. Wright, and A. Bittar, *Solid State Commun.* 59, 699 (1986).
83. M. W. Wright and H. J. Trodahl, *Phys. Rev. B* 36, 4383 (1987).
84. N. J. Shevchik, *J. Non-Cryst. Solids* 12, 141 (1973).
85. A. L. Greer, *Annu. Rev. Mater. Sci.* 17, 219 (1987).
86. K. Tanaka, Y. Ichimura, and M. Komasaki, *Thin Solid Films* 189, 51 (1990).
87. H. Shirai, S. Oda, T. Nakamura, and I. Shimizu, *Jpn. J. Appl. Phys.* 26, 991 (1987).
88. E. Vateva, R. Ionov, D. Nesheva, and D. Arsova, *Phys. Status Solidi A* 128, K23 (1991).
89. R. Ionov, D. Nesheva, and D. Arsova, *J. Non-Cryst. Solids* 137–138, 1151 (1991).
90. R. Ionov and T. Dudev, *Appl. Phys. A* 55, 203 (1992).
91. R. Ionov and D. Nesheva, *Solid State Commun.* 82, 959 (1992).
92. R. I. Ionov, *Europhys. Lett.* 19, 317 (1992).
93. R. Ionov and D. Nesheva, *Superlattices Microstruct.* 11, 439 (1992).
94. R. I. Ionov, *J. Non-Cryst. Solids* 159, 222 (1993).
95. E. Vateva and I. Georgieva, *J. Non-Cryst. Solids* 164–166, 865 (1993).
96. D. Nesheva, D. Arsova, and Z. Levi, *Philos. Mag. B* 70, 205 (1994).
97. E. Vateva and D. Nesheva, *J. Non-Cryst. Solids* 191, 205 (1995).
98. D. Nesheva, E. Vateva, Z. Levi, and D. Arsova, *Philos. Mag. B* 72, 67 (1995).
99. D. Nesheva, *Thin Solid Films* 280, 51 (1996).
100. E. Vateva, in "Physics and Applications on Non-Crystalline Semiconductors in Optoelectronics" (A. Andriesh and M. Bertolotti, Eds.), p. 61. Kluwer Academic, Dordrecht, 1997.
101. D. Nesheva, I. P. Kotsalas, C. Raptis, and E. Vateva, *J. Non-Cryst. Solids* 224, 283 (1998).
102. D. Nesheva, *J. Optoelectron. Adv. Mater.* 1, 13 (1999).
103. D. Nesheva, I. P. Kotsalas, and C. Raptis, *J. Appl. Phys.* 86, 4964 (1999).
104. D. Nesheva, *J. Optoelectron. Adv. Mater.* 3, 885 (2001).
105. S. Imamura, Y. Kanemitsu, M. Saito, and H. Sugimoto, *J. Non-Cryst. Solids* 114, 121 (1989).
106. H. Hamanaka, S. Konagai, K. Murayama, M. Yamaguchi, and K. Morigaki, *J. Non-Cryst. Solids* 198–200, 808 (1996).
107. A. Kikineshy, *Opt. Eng.* 34, 1040 (1995).
108. V. Palyok, A. Mishak, I. Szabo, D. L. Beke, and A. Kikineshi, *Appl. Phys. A* 68, 489 (1999).
109. A. Kikineshi, A. Mishak, V. Palyok, and M. Shiplyak, *Nanostruct. Mater.* 12, 417 (1999).
110. V. Palyok and M. Malyovanik, *J. Optoelectron. Adv. Mater.* 3, 77 (1999).
111. A. Csik, M. Malyovanik, J. Dorogovics, A. Kikineshi, D. L. Beke, I. A. Szabo, and G. Langer, *J. Optoelectron. Adv. Mater.* 3, 33 (2001).
112. A. Imre, V. Fedor, M. Kis-Varga, A. Mishak, and M. Shipljak, *Vacuum* 50, 507 (1998).
113. T. Wagner and P. J. S. Ewen, *J. Non-Cryst. Solids* 266–269, 979 (2000).
114. E.-K. Kim, S.-I. Kwun, S.-M. Lee, H. Seo, and J.-G. Yoon, *Appl. Phys. Lett.* 76, 3864 (2000).
115. M. Popescu, F. Sava, A. Lorinczi, E. Vateva, D. Nesheva, P.-J. Koch, T. Gutberlet, W. Uebach, and H. Bradaczek, *Solid State Commun.* 103, 431 (1997).
116. A. Jĕrōme and J. Commandeur, *Nature* 386, 589 (1997).
117. M. Arndt, R. Stannarius, H. Groothues, E. Hempel, and F. Kremer, *Phys. Rev. Lett.* 79, 2077 (1997).
118. L. Cristoflini, S. Arisi, and M. P. Fontana, *Phys. Rev. Lett.* 85, 4912 (2000).
119. H. Okamoto, K. Hattori, and Y. Hamakawa, *J. Non-Cryst. Solids* 164–166, 445 (1994).
120. K. Hattori, T. Mori, H. Okamoto, and Y. Hamakawa, *Phys. Rev. Lett.* 60, 825 (1988).
121. K. Saito, Y. Utsugi, and A. Yoshikawa, *J. Appl. Phys.* 63, 565 (1988).
122. Y. Utsugi, M. Kakuchi, and H. Maezawa, *Rev. Sci. Instrum.* 60, 2295 (1989).
123. Y. Somemura, A. Yoshikawa, and Y. Utsugi, *Jpn. J. Appl. Phys.* 31, 3712 (1992).

124. B. Singh, S. P. Beaumont, P. G. Bower, and C. D. W. Wilkinson, *Appl. Phys. Lett.* 41, 889 (1982).
125. G. H. Berstein, W. P. Liu, Y. N. Khawaja, M. N. Kozicki, D. K. Ferry, and L. Blum, *J. Vac. Sci. Technol., B* 6, 2298 (1988).
126. K. Mietzsch and A. G. Fitzgerald, *Appl. Surf. Sci.* 162–163, 464 (2000).
127. H. Y. Lee, S.-W. Paek, and H.-B. Chung, *Jpn. J. Appl. Phys.* 37, 6792 (1998).
128. Y. Isbi, S. Sternklar, E'l Granot, V. Lyubin, M. Klebanov, and A. Lewis, *Opt. Commun.* 171, 219 (1999).
129. T. Gotoh and K. Tanaka, *Proc. SPIE* 4088, 351 (2000).
130. T. Gotoh and K. Tanaka, *J. Appl. Phys.* 89, 4697 (2001).
131. R. Wiesendanger, *J. Vac. Sci. Technol. B* 12, 515 (1994).
132. Y. Utsugi, *Nature* 347, 747 (1990).
133. M. Ohto and K. Tanaka, *J. Non-Cryst. Solids* 227–230, 784 (1998).
134. M. Ohto and K. Tanaka, *J. Appl. Phys.* 92, 5468 (2002).
135. H. Hisakuni and K. Tanaka, *Science* 270, 974 (1995).
136. A. Hirotsune, Y. Miyauchi, and M. Terao, *Appl. Phys. Lett.* 66, 2312 (1995).
137. V. Weidenhof, I. Friedrich, S. Ziegler, and M. Wutting, *J. Appl. Phys.* 86, 5879 (1999).
138. V. Weidenhof, I. Friedrich, S. Ziegler, and M. Wutting, *J. Appl. Phys.* 89, 3168 (2001).
139. S. Hosaka, T. Shintani, M. Miyamoto, A. Hitotsune, M. Terao, M. Yosida, K. Fujita, and S. Kämmer, *Jpn. J. Appl. Phys.* 35, 443 (1996).
140. J. Tominaga, T. Nakano, and N. Atoda, *Appl. Phys. Lett.* 73, 2079 (1998).
141. T. Fukata, J. Tominaga, T. Nakano, and N. Atoda, *Appl. Phys. Lett.* 75, 3114 (1999).
142. S. R. Ovshinsky, *Romanian Rep. Phys.* 51, 171 (1999).
143. http://www.ovonyx.com.
144. K. Nakayama, K. Kojima, F. Hayakawa, Y. Imai, A. Kitagawa, and M. Suzuki, *Jpn. J. Appl. Phys.* 39, 6157 (2000).
145. H. Kado and T. Tohda, *Appl. Phys. Lett.* 66, 2961 (1995).
146. H. Kado and T. Tohda, *Jpn. J. Appl. Phys.* 36, 523 (1997).
147. T. Gotoh, K. Sugawara, and K. Tanaka, *J. Non-Cryst. Solids* 299–302, 968 (2002).
148. D. Saluel, J. Daval, B. Bechevet, C. Geermain, and B. Valon, *J. Magn. Magn. Mater.* 193, 488 (1999).
149. K. Sugawara, T. Gotoh, and K. Tanaka, *Appl. Phys. Lett.* 79, 1549 (2001).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanostructured Extracellular Matrix

Zuwei Ma, S. Ramakrishna

National University of Singapore, Singapore

CONTENTS

1. INTRODUCTION

In tissue engineering, cells are seeded into a biodegradable scaffold to obtain a cell-scaffold complex, which then will be implanted into patients to replace diseased or damaged tissue. In this process, the cells in the scaffold should be induced to proliferate and differentiate in a manner that mimics the natural process of tissue regeneration in the human body. To do so, the scaffold should be made an analog of the natural environment of the cells *in vivo*.

In human and animal bodies, tissues are not solely made up of cells. The insoluble noncellular material present between the cells throughout the tissue is known as the extracellular matrix (ECM), which is composed of various biomacromolecules secreted by the cells themselves. There are three major functions of the ECM. First of all, they provide structural support and tensile strength. Second, they provide substrates for cell adhesion and a highway for cell migration. Third, the ECM regulates cellular differentiation and metabolic functions, for example, by modulation of cell growth by the binding of growth factors. Not only the ECM composition, but also the three-dimensional organization of the biomacromolecules in ECM, of which the dimension is typically in the nanoscale, influence cellular behavior significantly. So, the mimic of the tissue-engineering scaffold to the natural ECM should be made not only in terms of the chemical composition, but also in terms of the 3D nano/microphysical structures. Here, the ECM analogs (scaffolds) are called synthetic ECM, to be distinguished from the natural ECM.

It is not difficult to prepare synthetic ECM with similar chemical compositions as natural ECM. For many years, biomacromolecules derived from ECM such as collagen [2], chitosan [3], hyaluronan [4], alginate [5], and so on, have been extracted, modified, and processed to prepare tissue-engineering scaffolds for the regeneration of skin, cartilage, liver, and many other human tissues. ECM-derived biomacromolecules or specific polypeptides have also been used to modify the surface of the synthetic biodegradable polymers such as PLLA, to obtain an ECM-mimic surface, while remaining the good mechanical properties of the bulk materials. For example, tripeptide RGDs, which can combine specifically with integrins on a cell membrane, have been introduced on the surface of many polymeric biomaterials to enhance cell adhesion [6].

Compared to the chemical modification of the synthetic ECM, fewer works have been done to prepare scaffolds with similar morphological characters with natural ECM. In addition, more attention has been paid to the influences of material chemical compositions on cell behavior than to the influences of material physical morphologies. The reason may be that there are many more difficulties in controlling material morphologies than in controlling chemical compositions. Scaffolds with precise micro- or nanostructures cannot be easily produced by some traditional technologies such as particulate leaching and the foaming method.

In this chapter, the relationships between the material morphologies and the cell behaviors in ECM and synthetic ECM will be reviewed. First, the chemical composition and the 3D morphology of ECM and the interactions between cells and ECM will be discussed to explain why nanoscale structures are so important to cell behavior. Then, nano- or microsurface patterns, which have been developed in recent years as a powerful tool for the fundamental research of interactions between surface morphologies and cells, will be discussed. Finally, nano- or microscale polymer fibers and nonwoven polymer meshworks used as tissue-engineering scaffolds will be introduced.

It should be noted here that many previous works have demonstrated that the dimensions of the physical structures

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (641–655)

that cells react to is about from 30 nm to 20 μm. Due to this reason, literature involving the structures with dimensions in this range will all be discussed in this review to provide a complete and comprehensive understanding, although the structures in natural ECM are mainly in the nanoscale.

2. COMPONENTS AND STRUCTURE OF EXTRACELLULAR MATRIX: NANOSCALE STRUCTURES

An extracellular matrix is composed of versatile biomacromolecules, of which there are three important classes. The first class is called glycosaminoglycans (GAGs), which belong to polysaccharide chains. The second class is fibrous proteins, such as collagen, elastin, and fibronectin. The third class is nonfibrous proteins, such as growth factors and cytokines.

Most of these biomacromolecules are secreted locally by the cells. Here, the synthesis and secretion of the collagen will be described briefly as an example. Genes encoding for the α chains of precollagen in DNA are transcripted to mRNA in the cell nucleus. The mRNA is translated by ribosomes on endoplasmic reticulum to synthesize the α chains. Modifications of the α chains such as hydroxylation of proline, hydroxylation of lysine, and glycosylation of some hydroxlysine are made, and the α chains assemble into triple helix precollagen molecules in golgis. Finally, the procollagen molecules are secreted by exocytosis into the ECM space. In ECM, "registration peptides" in precollagen molecules are cut down by precollagen peptidases, resulting in collagen molecules, which are one of the most important proteins in ECM.

2.1. GAGs

GAGs are unbranched polysaccharide chains composed of repeating disaccharide units. One of the two sugar residues in the repeating disaccharide is always an amino sugar (*N*–acetylglucosamine or *N*–acetylgalactosamine), which in most cases is sulfated. The second sugar is usually an uronic acid (glucuronic or iduronic). Four main groups of GAGs most frequently found in an ECM are: (1) *hyaluronan*, (2) *chondroitin sulfate* and *dermatan sulfate*, (3) *heparan sulfate* and *heparin*, and (4) *keratan sulfate*. In most cases in an ECM, large numbers of GAG chains are covalently bonded with a protein core to form proteoglycan, called aggrecan (Fig. 1 [7]). The molecular weight of the protein core is only a very small fraction of the total molecular weight of the proteoglycan. It should be noted that hyaluronan does not bind with proteins to form proteoglycans.

Due to the plentiful sulfate or carboxyl groups in GAG chains, the proteoglycan molecules are highly negatively charged, and combine large amounts of water to form gel-like hydrogels. In the ECM of articulate cartilage, proteoglycan molecules are noncovalently connected with hyaluronan to form huge polysaccharide aggregates (Fig. 1 [7]). The strongly hydrophilic aggregates combine huge amounts of water to form hydrogels, in which the GAG chains adopt highly extended conformations that occupy a huge volume relative to their mass. Moreover, the high density of negative charges in GAG chains attract a cloud of cationic ions, such as Na^+, that are osmotically active, causing large amounts of water to be sucked into the matrix. This creates a swelling pressure that enables the matrix to withstand very strong compressive forces, which is one of the most important functions of the articulate cartilage.

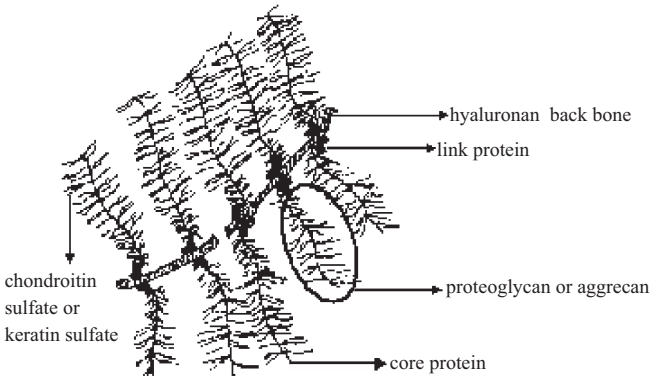


Figure 1. A huge polysaccharide molecule in the ECM of hyaline cartilage. Reprinted with permission from [7], B. L. Seal et al., *Mater. Sci. Eng.* R34, 147 (2001). © 2001, Elsevier Science.

2.2. Fibrous and Nonfibrous Proteins

The proteins in an ECM come in a great variety of shapes and sizes. Most of these proteins are in fibrous forms, while a small part are in nonfibrous forms. There are two functional types of fibrous proteins: the ones that are mainly structural, like collagen and elastin, and the ones that are mainly adhesive, like fibronectin and laminin. Growth factors in an ECM, which have a crucial effect on the cell growth and differentiation, belong to nonfibrous proteins.

The most important fibrous protein in human and animal ECM is collagen. About 15 types of collagen molecules have been found. The main types of collagen are type I, II, III, V, and XI. Type II collagen is the principal collagen in cartilage tissue, while type I collagen is by far the most common, and is the principal collagen of skin and bone. The collagen fiber network in an ECM provides the tensile strength for the tissue. Another important structural protein is elastin fibers, which provide the elasticity for the tissue. Fibronectin and laminin belong to the adhesion proteins, and are also in a fibrous form. The fibronectin or laminin fibers have specific biding sites, either for other matrix biomacromolecules or for receptors on the cell surface. These protein fibers thus contribute to both organizing the biomacromolecules in an ECM and helping attach the cells to the ECM. Recent research has demonstrated that collagen also has cell-binding sites in their molecules [8]. So collagen belongs to both structural and adhesion proteins. Table 1 summarizes some main components of the extracellular matrix and their function.

2.3. 3D Organization of Biomacromolecules in ECM

Once secreted into the ECM, the biomacromolecules will be orientated and assembled into an organized 3D meshwork, either by cell orientations or by self-assembly. In different

Table 1. Main components of the extracellular matrix and their function.

Components	Structure	Function
Collagen	Triple helical glycoprotein. Molecules are mainly composed of proline, hydroxyproline, and glycine.	Provides tensile strength to tissues; support and structure for tissues and organs; mediates cell adhesion to ECM.
Elastin	Stretchable hydrophobic protein interacting with glycosylated microfibrils.	Provides tissue with elasticity; allows for expansion and contraction of the tissue.
Fibronectin	Specialized adhesive glycoprotein.	Mediates cell–ECM adhesion by binding with cells and other biomacromolecules in ECM.
Laminin	Large, complex adhesive glycoprotein.	Mediates cell–ECM adhesion in connective tissue; binds cells to type IV collagen and heparin sulfate in basal laminae.
Proteoglycans	Heterogeneous, long glycosaminoglycans (GAG) chains covalently linked to a core protein.	Combines with large amounts of water to form hydrogels, shock absorption, sequestration of growth factors, and cytokines.
Hyaluronic acid	A very large, specialized, nonsulfated GAG. Does not form proteoglycan.	Provides a fluid environment for cell movement and differentiation; binds the growth factors and cytokines.
Growth factors and cytokines	Protein molecules, embedded in ECM. Combined with protein fibers or proteoglycan hydrogels by noncovalent interactions.	Regulation of cell functions and gene expressions.

tissues, variations in the relative amounts of the different types of biomacromolecules and the way they are organized give rise to an amazing diversity of ECM morphologies, each adapted to different particular functional requirements.

However, there are some common characteristics in the structure of various ECMs. Here, connective tissue can be used to describe the 3D structure of an ECM. In connective tissue, such as skin, cartilage, liver, tendons, and corneas, there is many more plentiful ECMs than cells. In an ECM of connective tissues, both the fibrous and nonfibrous proteins are embedded in the large volume of hydrogels composed of proteoglycan and waters. The fibrous proteins, such as collagen, elastin, fibronectin, and laminin fibers, penetrate through the hydrogels to form a 3D network (Fig. 2 [9]). The growth factors, cytokines, and other nonfibrous proteins are dispersed in the hydrogel or combined with the fibrous proteins.

The discussion of the ECM structure above has focused on the volume-filling ECM between cells. But in certain places, especially beneath epithelia, the ECM can also be organized as a thin tough sheet–basal lamina, or basement membrane. Basal laminae are flexible, thin (40–120 nm thick) mats of specialized ECM that underlie all epithelial cell sheets to separate them from the underlying connective tissues. Most mature basal laminae contain mainly type IV collagen, laminin, and the large heparan sulfate proteoglycan, perlecan. The type IV collagen and laminin form a dense mesh, which is embedded in the perlecan hydrogel. In some cases, such as in kidney glomerulus, the basal lamina lies between two epithelial cell sheets, and functions as a highly selective filter [10]. Basal laminae serve in more than simple supporting and filtering roles. They influence cell metabolism, organize the proteins in adjacent cell membranes, induce cell differentiation, and serve as specific highways for cell migration. Figures 3 [11], 4 [12], and 5 are schematic diagrams of the structures and locations of basal laminae.

2.4. Nanoscale Structure of ECM

Until now, the diameter of the fibers embedded in the proteoglycan hydrogel in an ECM have not been mentioned. It has been pointed out that the most common protein fiber in an ECM is collagen fiber. The characteristic feature of a typical collagen molecule is its long, stiff, triple-stranded helical structure, in which three collagen polypeptide chains, called α chains, are wound together to form a rope-like superhelix, the collagen molecule. After being secreted into the extracellular space, the collagen molecules assemble into ordered polymers called collagen fibrils (10–300 nm in diameter, Fig. 6), which can aggregate further into larger, cable-like collagen fibers with diameters of several hundred nanometers or several microns (Fig. 7 [13, 14]). Nishida et al. [15] observed the protein matrix composed largely of collagen fibers in the cornea of a rat after the hydrogel in the rat cornea was removed by enzyme and acid treatment.

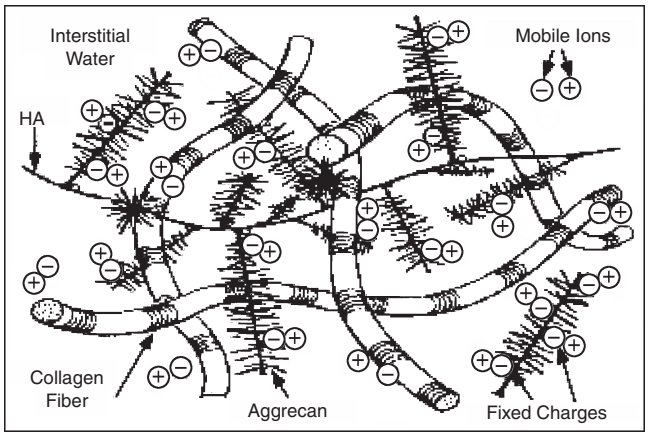


Figure 2. Schematic of an ECM in a hydrated connective tissue. Reprinted with permission from [9], V. C. Mow et al., *Int. J. Solids Struct.* 35, 4945 (1998). © 1998, Elsevier Science.

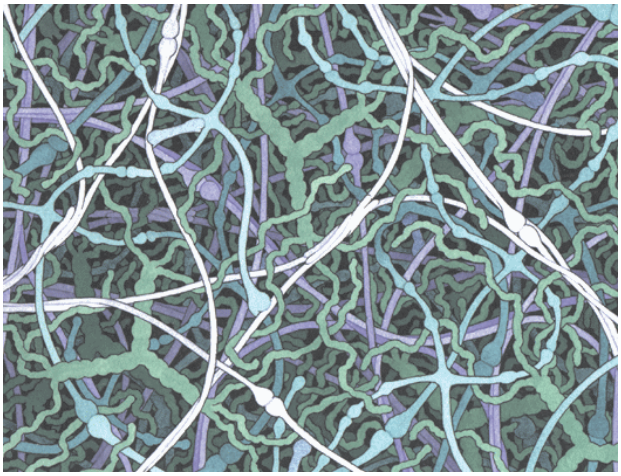


Figure 3. Illustration of a basement membrane, which forms a tough surface that supports the skin and many organs. Collagen type IV forms the structural basis. Collagen type IV fibrils entangled with two other molecules, laminin, and proteoglycans to form a dense sheet. Reprinted with permission from [11], http://www.rcsb.org/pdb/molecules/pdb4_3.html. © David Goodsell.

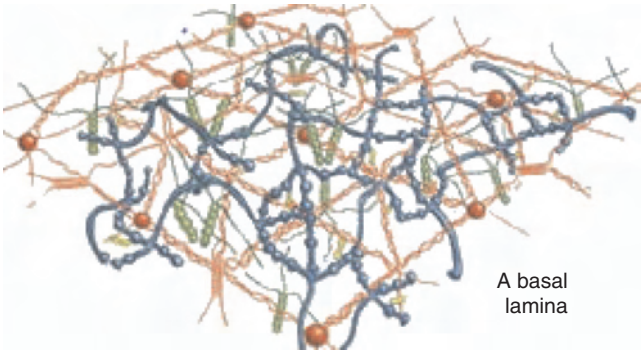


Figure 4. The basal lamina is a network of extracellular molecules. It can serve as a mechanical support and as a sieve that will exclude proteins, but allow small molecules to pass through it. Reprinted with permission from [12], http://www.virtuallaboratory.net/BioFunKFI/lectureNotes/Topic6-3_Adhesion.htm. © M. W. Klymkowsky.

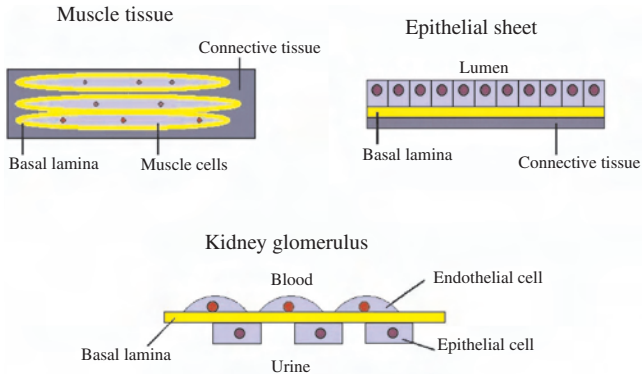


Figure 5. Locations of basal lamina.

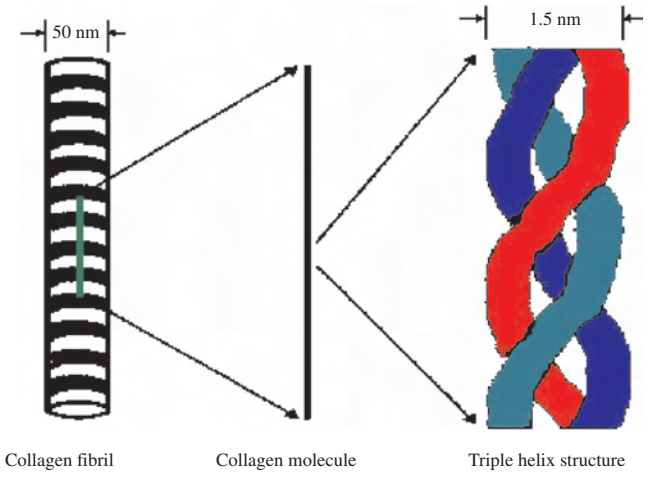


Figure 6. Schematic representations of a collagen molecule and a collagen fibril.

The diameter of collagen fibers in a rat cornea was about 100–400 nm.

The fiber diameters in the basal laminae are also in the nanometric range. Research has shown that basal laminae are comprised of a complex mixture of pores, ridges, and fibers [16]. Shirato et al. [17] studied the glomerular basement membrane of the rat kidney, in which the fibrils are 5–11 nm thick and the pores are 10–30 nm wide. Yamasaki and co-workers examined the pores and fibers in bovine glomerular and tubular basement membranes [18]. The pore diameter is about 10–12 nm, and the fiber diameter is 3–15 nm. Another study by Hironaka et al. found that the glomerular, tubular, and Bowman's capsule basement membranes of the rat kidney all had mean fiber diameters of 6–7 nm and mean pore diameters of 10–13 nm [19]. Using SEM, TEM, and AFM, Abrams et al. [20] measured the feature sizes on the surface of corneal epithelial basal laminae of the Macaque monkey. The average feature heights were from 147 to 191 nm, and the average fiber width was 77 nm. They found that pores made up 15% of the total surface area of the membrane, and had an average diameter of 72 nm. Similar features were observed on the human corneal epithelial basement membrane and on Matrigel™, a commercially available basement membrane matrix [21].

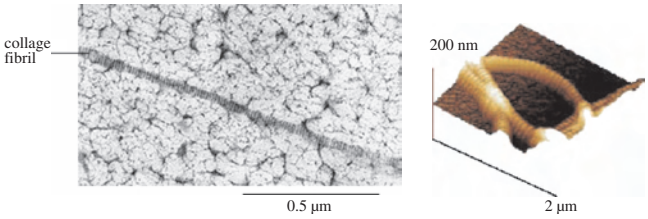


Figure 7. Left: TEM picture of a collagen fibril. Reprinted with permission from [14], E. B. Hunziker et al., *J. Cell Biol.* 98, 277 (1984). © 1984, Rockefeller University Press. Right: AFM picture of collagen fibrils and fibers. Reprinted with permission from [13], S. Mosler et al., "Symposium on Nano-Physics of Life Systems," Copenhagen, June 2002, poster. © 2002, S. Mosler and N. B. Larsen.

A scanning electron micrograph of a corneal epithelial basement membrane is shown in Figure 8 [16].

Now, it can be concluded that, in a natural ECM, protein fibrils and fibers with mainly nanometric-scale diameters entangle with each other to form a nonwoven meshwork, which is embedded in proteoglycan hydrogel. In basal laminae, the meshwork is much denser and the fibers much thinner than in other ECMs in common connective tissue. Due to this characteristic of the ECM, in recent years, much interest has been focused on the fabrication of biodegradable polymeric nanofibers and nonwoven meshworks to be used in the field of tissue engineering and artificial organs [22].

3. CELL INTERACTIONS WITH EXTRACELLULAR MATRIX AND BIOMATERIALS

The interactions between cells and their ECMs are mediated by a process called biorecognition. In this process, the transmembrane proteins called receptors on the cell membrane combine specifically with the corresponding ligands in the ECM, causing a series of events in the signal transduction cascade within the cells, and eventually influencing their gene expression. For example, growth factors such as FGF combine with their receptors on the cell surfaces, and stimulate their proliferation and differentiation.

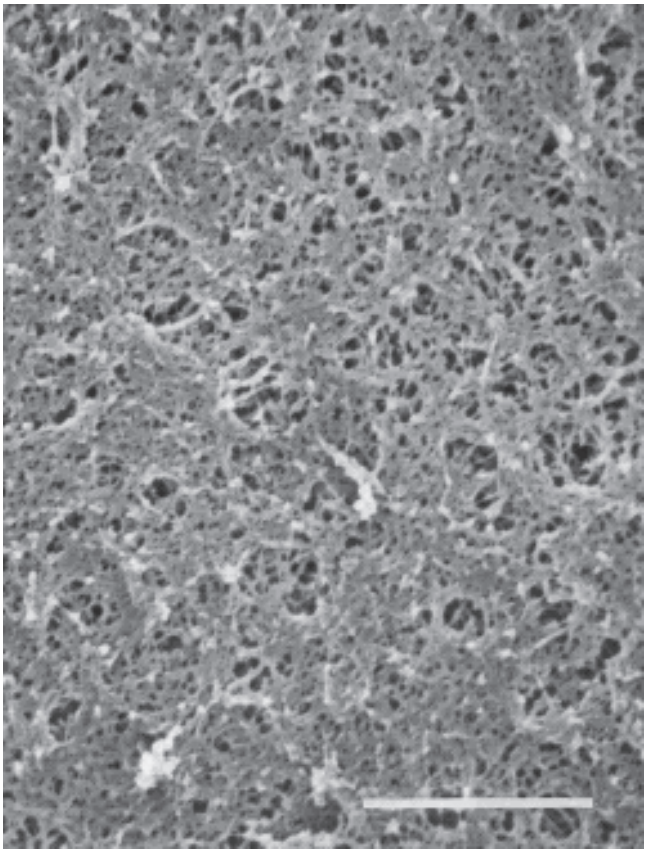


Figure 8. Scanning electron micrograph of a corneal epithelial basement membrane of Macaque monkey. After the corneal epithelium was stripped, the cornea was fixed in 2% glutaraldehyde, dried by the critical point method, and imaged on a Hitachi S-900 scanning electron microscope (bar=1 m). Reprinted with permission from [16], R. G. Flemming et al., *Biomater.* 20, 573 (1999). © 1999, Elsevier Science.

3.1. Integrins

Many kinds of receptors exist on a cell membrane, among which the integrins are the most important. The interactions between the integrins and their ligands in an ECM are the main ways that cells bind to and respond to the ECM. This process is involved in a variety of cellular functions such as wound healing, cell differentiation, and apoptosis.

Integrins belong to a large family of cell adhesion receptors involved in cell–ECM and cell–cell interactions. Functional integrins consist of two transmembrane glycoprotein subunits, called α and β, which are noncovalently bound (Fig. 9 [23]). Both of the subunits contribute to the binding of the ligands. Until now, 16 alpha and 8 beta subunits have been identified. From these subunits, only 22 integrins are formed in nature, which implies that not all possible combinations exist. The ligands of integrin are mainly adhesion proteins, including fibronectin, laminin, vitronectin, and collagens. The biding sites for integrins in adhesion proteins are relatively small, typically five amino acid residues in length or less. The first integrin-binding site to be identified was the tripeptide RGD (arginine–glycine–aspartic acid). Figure 10 [24] is a schematic diagram illustrating the interactions between integrins and fibronectin fibers in an ECM.

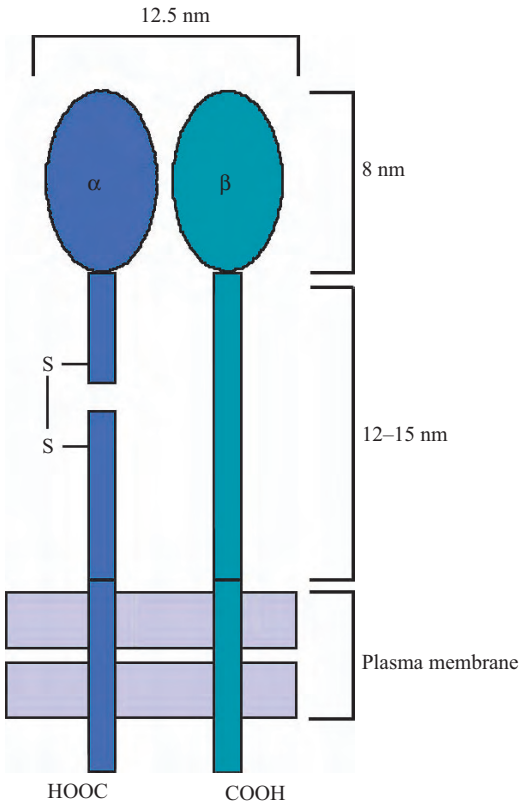


Figure 9. Integrin structures.

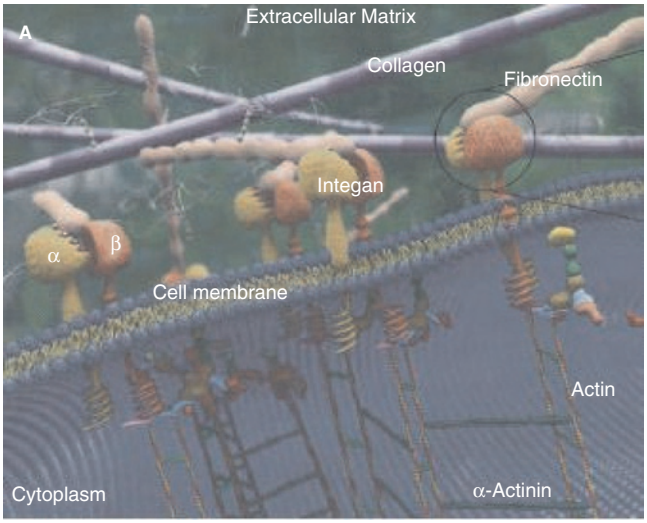


Figure 10. Interactions of integrins and their ligand, adhesion proteins in ECM. Reprinted with permission from [24], M. Tirrell et al., *Surf. Sci.* 500, 61 (2002), © 2002, Elsevier Science.

Another adhesion receptor important for cell–ECM adhesion is cell surface proteoglycans, which are present on cell surface and serve as adhesion receptors with a lower specificity than integrins. The negative charges of GAG chains in the cell surface proteoglycan can interact with regions of clustered positive charges on adhesion proteins such as fibronectin. Other key adhesion receptors on cell surfaces include cadherins, selectins, and Ig superfamilies like NCAMs (neural cell adhesion molecules) and ICAMs (intercellular adhesion molecules). They are all transmembrane proteins, and are important to cell–cell junctions. Cadherin on a cell surface can bind another cadherin on another cell. Selectin on a cell surface can bind some specific oligosaccharides on another cell surface. NCAM can bind two nerve cells through interactions between two NCAMs. Integrins are also involved in cell–cell interactions through the binding of an integrin on a cell surface to its ligand on another cell surface.

3.2. Focal Contact

Focal contact, also called focal adhesion, is formed when a cluster of integrins accumulate in a relatively small spot on the cell membrane and interact with adhesion proteins in an ECM. Focal contact is the main way in which cells attach onto the ECM. The extracellular part of the integrin combines with the adhesion proteins in ECM, while the intracellular part of the integrin combines with the cytoskeleton (actin filament) through other cytoskeletal proteins such as α-actin, vinculin, and talin. By doing so, the cytoskeleton of a cell is attached physically to the ECM. However, one integrin cannot bind its ligand strongly. This is because the affinity between the integrin and its ligand is much lower than that of other cell-surface receptors. When the integrins are evenly distributed over the cell surface, no adhesion will occur. But the integrins are usually present at 10–100-fold higher concentration on the cell surface than other receptors. In an integrins-clustering focal contact, the combined weak affinities can give rise to a spot on the cell surface that has enough adhesive (sticking) capacity to adhere to the extracellular matrix (Fig. 11 [25]). In this way, cells can bind simultaneously, but weakly to large numbers of matrix molecules, and still have the opportunity to explore their environment without losing all attachment to it. If the receptors were to bind too strongly to their ligands, cells would probably be irreversibly bound to the matrix, depriving it of motility.

Through the formation of focal contacts, dynamic sites are built to mechanically bind the cytoskeleton to the ECM. The mechanical signal can be transferred into the cytoskeleton, and sensed by some stretch receptors that can help to reorganize the cytoskeleton structure, and eventually regulate the shape and migration of the cell. Produced by the clustering and activation of the integrins on the cell surface, focal contact also mediates the transduction of other extracellular signals into the intracellular signal transduction pathways. Focal adhesion kinase (FAK), mitogen-activated protein kinase (MAPK), and integrin-linked kinase (ILK) are often involved in the signal transduction mediated by the focal contact (Fig. 12).

FAK is a tyrosine kinase commonly found in integrin-mediated focal adhesions. Upon activation and phosphorylation of FAK, this kinase may phosphorylate other signaling proteins in a signal transduction cascade. Paxillin, for example, a protein involved in cytoskeletal reorganization, is a target of FAK. The activation of MAPK occurs after integrin binding RGD peptides in fibronectin or laminin, resulting in the translocation of MAPK from the cytoplasm to the nucleus. MAPK can also be activated by integrin-linked kinase (ILK) in a FAK-independent pathway.

3.3. Why Nanometric Scale Structure?

From Figure 9, it can be seen that the diameter of one integrin is about 10–15 nm, while the focal contact will be much larger. No works have been done to study how many integrins there are in a focal contact, probably because that interactions between integrins and their ligands is in a dynamic equilibrium. However, some works mentioned that the diameter of a focal contact is about 0.5–2 μm in width and 5–10 μm in length [26, 27]. Another work pointed that the dimension of an individual focal adhesion is about 3–5 μm [28]. It can be hypothesized that the dimension of the protein fibers in the ECM, which is typically from several

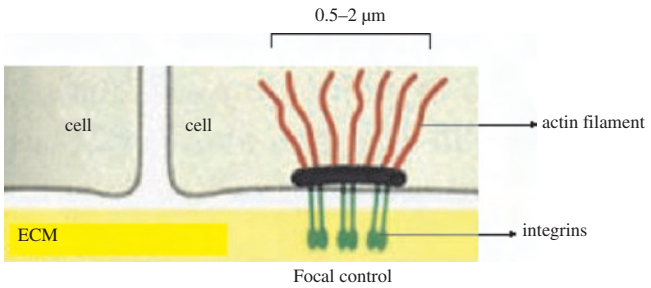


Figure 11. Focal contact formation. Reprinted with permission from [25], http://www.virtuallaboratory.net/BioFun-KFI/lectureNotes/Topic6-3_Adhesion.htm. © M. W. Klymkowsky.

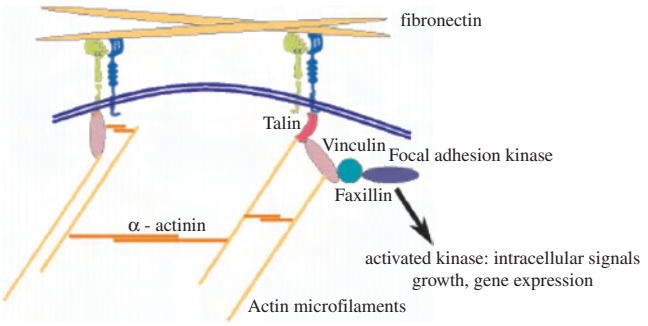


Figure 12. Integrin-mediated cell adhesion and signal transduction.

ten to several hundred nanometers, is the minimum for the formation of a focal contact. The focal contact formed on a protein fiber can extend along the fiber, and can be several micrometers in length. This feature may be very important and useful. The cells may be very easily induced to adhere, spread, and orient along the nanoscale fibers. Due to this kind of induction, protein fiber networks in an ECM can act as highways for the migration of the cells.

3.4. Cell Interactions with Biomaterial Surface

Cells interact with biomaterials in the same way as they interact with the ECM *in vivo*. Some adhesion proteins such as fibronectin, laminin, and vitronectin are soluble in blood plasma and other body fluids. The adhesion proteins also can be found in the cell culture media in which fetal bovine serum is frequently added. Upon contacting the biomaterials, the adhesion proteins are spontaneously adsorbed on the biomaterial surface, and act as ligands for cell integrins. As in an ECM, the cells also adhere on the biomaterial surface through the formation of focal contacts.

4. NANO- AND MICROSURFACE PATTERNS

Both the chemical compositions and physical structures in the ECM have great influences on the cell behavior, the mechanism of which is far from being well understood. Except for the adhesion proteins, many other ligands, such as hyaluronan and growth factors, exist in an ECM. They can also bind the receptors on the cell surface to adjust cell functions. Moreover, the physical morphologies of an ECM have great variety. Protein fiber meshworks with different nano- or microtextures in the ECM have different effects on cell adhesion, orientation, and migration. All of these chemical and physical factors entangle together, often synergistically, to affect cell behavior. In this process, so many factors are involved that it is impossible to study the effect of a single factor *in-situ* in an ECM because one cannot change one factor without altering the others.

However, using plane surfaces with specific chemical or physical properties, the research work can be greatly simplified. The cells can be cultured on the surfaces *in vitro*, and the adhesion, proliferation, migration, and differentiation can be observed. Thousands of experiments have been done using the plane surface as a research model in the last 20 years. The results showed that surface energy [29, 30], hydrophilicity [31, 32], surface charge [33–35], functional groups [36, 37], surface roughness [30, 38, 39], and surface topographical properties [40] affect the cell behavior significantly, and can be used to control the bioreactions on biomaterial surfaces. For example, immobilizing the adhesive proteins or peptides on a biomaterial surface has already been a traditional method to enhance cell adhesion on biomaterial surfaces [6].

In recent years, surfaces with defined chemical or physical patterns have been used as a much more powerful tool than homogeneous surfaces to study the relationships between the chemical or physical properties and the cell behavior. Much more information can be obtained using patterned surfaces compared to using homogeneous surfaces. The techniques for producing surface patterns are not new, however. Microfabrication technologies, which have been widely used in the microelectronics industry, are now gaining more and more attention in biomedical studies. These technologies allow the production of defined physical and chemical features on different substrates. The feature size, shape, and distribution on the surfaces can be precisely controlled and systematically changed. Photolithography and microcontact printing methods are the two most frequently used technologies to prepare surface patterns on biomaterial surfaces.

In this part, the preparation of the micro- or nanoscale surface patterns and their effect on cell functions will be introduced. Here, all kinds of surface patterns, no matter whether they are topographical or chemical, and no matter whether they are in the nanoscale or microscale, will be discussed.

4.1. Random Roughness

Almost any material has a surface with random roughness. For fundamental research, materials with different degrees of roughness can be made by sandblasting or a chemical/physical etching treatment. Lampin et al. [38] prepared a PMAA surface with different degrees of roughness (average roughness, $Ra = 0.07$–$3.3\ \mu$m) by a sandblasting treatment with alumina grains. Cell culture experiments showed that the adhesion and migration of cells on the PMMA surface increased with the degree of roughness. Similar results were also reported by Hallab [30] and Deligianni [39]. Fan et al. [41] prepared an Si wafer with different degrees of nanoscale roughness ($Ra = 2$–100 nm) using an HF or HF/HNO_3 etching treatment. More neural cells adhered on the surfaces with Ra ranging from 20 to 50 nm than on the surfaces with Ra less than 10 nm or higher than 70 nm. The viability of the cells was affected in a similar manner. The neural cells have higher tyrosine hydroxylase activity on the surfaces with Ra ranging from 20 to 50 nm than on other surfaces.

4.2. Regular Patterns

UV photolithography, which is widely used in the microelectronics industry to produce micropatterns on silicon wafer, is now being used to prepare morphological patterns on different material surfaces. The process of photolithography is schematically diagrammed in Figure 13. Surface-patterned

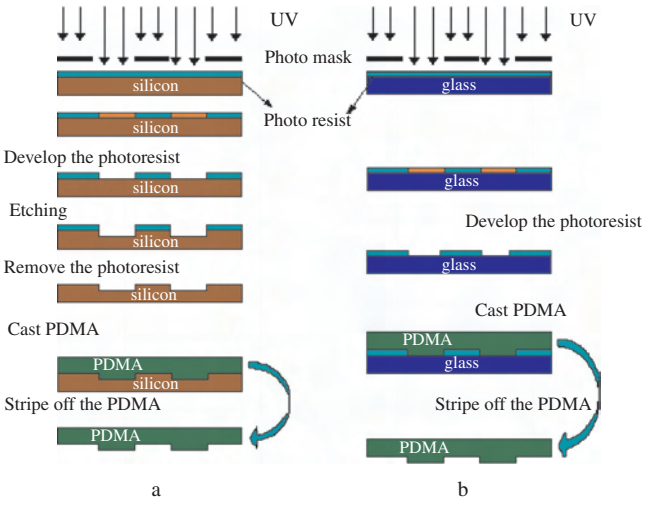


Figure 13. Photolithography technology to prepare surface-patterned PDMA. The PDMA stamp can be used as a master to prepare other surface-patterned materials.

silicon wafers are often produced first to be used as masters to replicate the surface pattern on other materials. In Figure 13(a), a photoresist-coated silicon wafer is exposed to UV light through a photomask. Then the photoresist is developed and the wafer is subjected to the etching, either by chemical erosion or by ion or electron-beam treatment. The depth of the pattern can be controlled by changing the etching time. Upon removing the remaining photoresist, topographical patterns can be produced on the silicon wafer. The surface-patterned silicon wafer can be used as a master to make polydimethylsiloxane elastomer (PDMS) duplicates. The surface-patterned PDMS has excellent thermal and chemical stability, and can be subsequently used as a master to produce topographical or chemical patterns on other material surfaces [42]. To cast a replica, the desired polymer solution is allowed to dry on the top of the PDMS master, and then the polymer film is stripped off and used. The PDMS can also be pressed into the polymeric material which has been heated to soften it. In Figure 13(b), no silicon wafer is used; the photoresist is directly coated on a glass substrate, irradiated by UV light, and developed. The macromonomer of PDMS is then cast on it and cured to obtain PDMS masters. Although many kinds of surface-patterned polymers can be made by photolithography, the most commonly used materials in research work are silicon, PDMS, or photoresist material.

Grooves and ridges (Fig. 14) are among the most studied topographical patterns related to cell morphology control, of which very excellent reviews have been made by Curtis, Wilkinson, and Flemming [40, 16]. Microgroove/ridge surfaces have shown significant control over cellular behavior, of which the most important phenomenon is that the cell spreading, alignment, and migration can be obviously oriented along the grooves/ridges (Figs. 15 [43] and 16 [44]). Wojciak-Stothard et al. [45] reported that murine P688DI macrophages were well oriented along the multiple grooves/ridges (0.5, 5 μm in depth, and 10 μm in width) on silica substrata. Moreover, they migrated mainly along the groove/ridge edges.

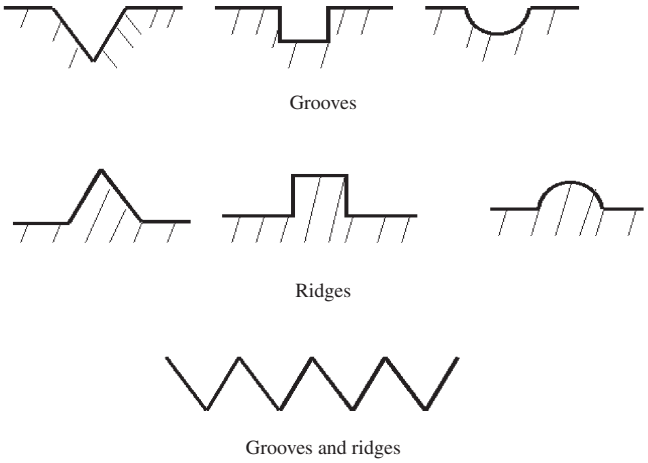


Figure 14. Grooves and ridges.

One theory that accounts for the cell orientation on the topographical surface is called the "contact guide effect." Integrins in focal contact transfer the variable degrees of tension or compression into the cytoskeleton. Cell stretch receptors subject to these stresses will be activated and reorganize the cytoskeleton according to the surface topography [40]. Curtis et al. put forward the idea that cells react to discontinuities. Focal contacts [46], vinculin [44], and F-actin [44, 45] are found to align over or close to the groove/ridge contact. Another theory [47] is that the changes of surface free energy due to the edges and disruptions may be the reason for the cell orientation.

The cellular alignment along the surface topographies is dependent on the dimensions of the surface features, a complex situation, varying with the cell types, cell–cell interaction, as well as substrate composition and topography. Clark et al. [48] reported that BHK cells and chick embryonic neural cells react to a single step, and the extent of the reaction increases with the increasing step height in the range of 1–20 μm. In another work, Clark et al. [49] cultured three cell types (BHK, MDCK, and chick embryo cerebral neurons) on multiple microgroove surfaces with varying height and width (4–24 μm repeat, 0.2–1.9 μm depth). It was found that repeat spacing (width) had a small effect (alignment is

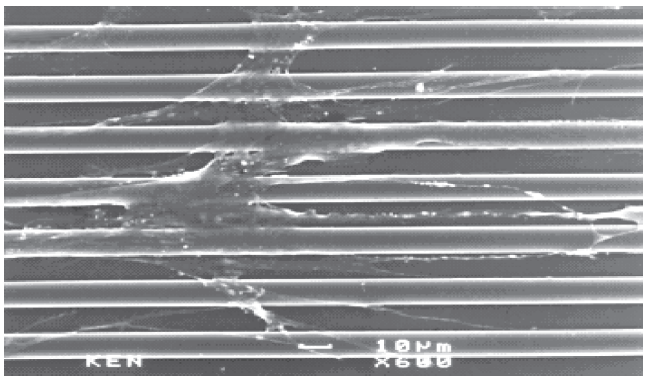


Figure 15. Osteoblasts oriented on microgrooved PS surface. Reprinted with permission from [43], K. Matsuzaka et al., *Biomater.* 20, 1293 (1999). © 1999, Elsevier Science.

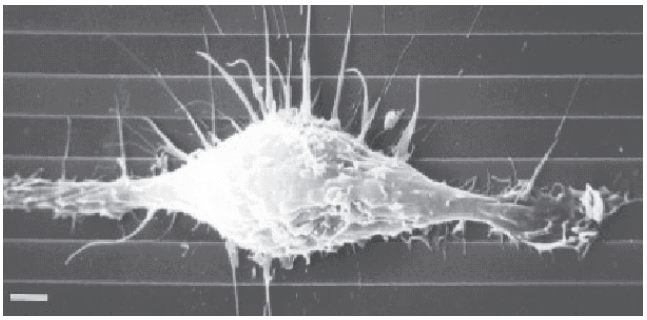


Figure 16. Well-spread macrophage with numerous microvilli on microgrooved substratum, groove depth 71 nm (bar is 5 μm). Reprinted with permission from [44], B. Wojciak-Stothard et al., *Exp. Cell Res.* 223, 426 (1996). © 1996, Elsevier Science.

inversely proportional to the spacing), but that the groove depth was much more important in determining cell alignment, which increased with the groove depth. Curtis and Wilkinson pointed out that, generally, the extent of cell orientation increases with groove depth from topographies of about 1 μm relief up to about 25 μm relief [40]. However, different results have also been reported. Wojciak-Stothard et al. [45] studied the orientation and migration of murine P688DI macrophages on a silica surface with microgrooves/ridges (0.5, 5 μm in depth and 10 μm in width). The results showed that cells speared and oriented better on the shallower (0.5 μm in depth) grooves/ridges. The macrophage migration on the shallower groove (0.5 μm in depth) was slightly faster, and had a longer persistent time than on the deeper grooves (5 μm in depth).

The smallest depth that a cell can sense has been studied, and found to be related to cell types. Clark et al. [50, 51] studied the elongation of BHK cells, MDCK cells, and chick embryo cerebral neurones on 260 nm period grating surfaces (130 nm grooves separated by 130 nm) of various depths (100, 210, and 400 nm). Only the BHK cells are aligned precisely by the grooves, while the other two kinds of cells were unaffected by the gratings. Wojciak-Stothard et al. [44] discovered that P388D1 macrophage-like cells could sense dimensions as small as 70 nm (Fig. 16). Casey et al. [52] reported that biological cells respond to topographic features in their environment at least as small as 30 nm. Such small topographies approach the dimensions of larger molecules, which are typically from several to several tens of nanometers. The groove width also affects the cell orientation degree [40]. When the grooves or ridges are appreciably wider than the cells, the effects on orientation are not very marked, although a cell may align to an edge. As the groove/ridge width is reduced to the width of the cells or less, the effects on cell orientation become more marked. No work has yet discovered the minimum width of topography to which a cell can react.

Other cell behavior such as cell adhesion and differentiation is also affected by groove/ridge patterns. Using UV photolithography and reactive ion etching, Wilkerson et al. [42] prepared rectangular ridges on the silicon wafer. Polystyrene solution was cast on the silicon wafer master to obtain a negative replication, and then PDMS was cast and polymerized on the polystyrene master to obtain a positive replication on the PDMS surface. On the surface-patterned PDMS, the 5 μm wide ridges (1–5 μm high) enhanced the adhesion of a kind of marine spore, the *Enteromorpha* spore, while they decreased the adhesion of endothelial cells. Matsuzaka et al. [43] prepared micropatterns with groove depths of 0.5, 1.0, and 1.5 μm and widths of 1, 2, 5, and 10 μm on PLA or PS surfaces by directly casting the polymer solution on micropatterned silicon wafers. Rat bone marrow cells were cultured on the microgrooved surfaces. Alignment of the cells along the grooves was observed (Fig. 15), and more mineralized ECM was formed on the microgrooved surfaces than on the smooth ones. PLA surfaces with a groove depth of 1.0 μm and a width of 1 and 2 μm induced the most mineralized ECM. The alkaline phosphatase activity of the cells was also higher on the microgrooved surfaces compared to the smooth ones.

Except for grooves and ridges, other topography was also prepared. Mata et al. [53] produced micropillars 6 μm in height and 5, 10, 20, and 40 μm in diameter on the surface of PDMS. Human bone-marrow-derived connective tissue progenitor cells (CTPs) spread, and adhered more on the micropillared PDMS than on the smooth PDMS. Tan et al. [54] fabricated polyimide (a kind of photoresist) pillars on glass and holes in polyimide film using UV photolithography. Arrays of the pillars and the holes with 2 μm, 4 μm height (or depth) separated by 10 μm had obvious effects on the cell spreading and migration. Neutrophils attached and spread on the smooth glass surface and surfaces with pillars. In contrast, cells were rounded, and did not adhere to either smooth polyimide film or microholed polyimide films. The migration of neutrophils was much faster on holes than on the polyimide surface, but it was significantly slower on pillars than on glass.

In addition to physical topography patterns, chemical patterns were fabricated to affect cell reactions. Mircocontact printing (μCP) is widely used in fabricating chemical patterns on biomaterial surfaces. This technique is outlined schematically in Figure 17 [55]. A PDMS stamp with the desired physical surface patterns is coated with the solution containing the substance that will be printed, such as adhesion protein. By contacting the stamp and the substrate, the substance can be printed on the substrate with the same patterns of the PDMS template. Physical interactions such as the strong adsorption between proteins and hydrophobic surfaces or chemical reactions such as that between an —SH group and gold surface can be used in μCP to maintain the stability of the micropatterned surfaces.

The patterning of adhesion proteins using μCP on biomaterial surfaces is an effective approach to control cell adhesion, spreading, orientation, and to obtain cell mircopatterns. Singhvi et al. [56] imprinted the hydrophobic compound $HS(CH_2)_{15}CH_3$ on a gold surface with a micropatterned PDMS stamp, then hydrophilic thiol compound $HS(CH_2)_{11}(OCH_2)_6OH$ was reacted with the remaining gold surface. The selective adsorption of laminin on the hydrophobic domain can be used to obtain rectangular islands of laminin with different sizes (40/15, 60/15, 70/20, 80/20 μm) to control the cell extension degree. DNA synthesis of hepatocytes increased greatly with the cell extension degree, while the albumin secreted into the medium decreased with the cell extension degree.

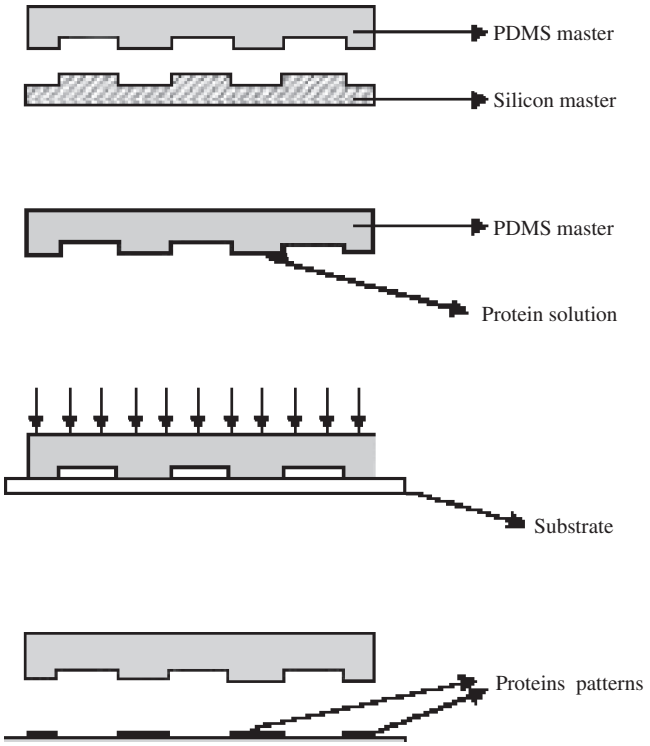


Figure 17. Schematic presentation of protein patterning process using microcontact printing technology. Reprinted with permission from [55], L. Kam et al., *Biomater.* 22, 1049 (2001). © 2001, Elsevier Science.

Fibronectin islands were also imprinted on the gold surface using the same method, and similar results on endothelia cells were obtained [28]. Kam et al. [55] prepared a honeycomb-like hexagonal micropattern of polylysine-conjugated laminin on a borosilicate glass surface using μCP technology. The hippocampal neurons selectively adhered to the polylysine-conjugated laminin area, and could be guided with high precision by polylysine-conjugated laminin patterns as narrow as 2.6 μm in width.

The self assembly monolayer (SAM) is another important technology to prepare surfaces with precise chemical or physical patterns. Self-assembly is the autonomous organization of components into patterns or structures without human intervention [57]. A SAM of amphiphilic molecules formed by hydrophobic interactions as in phospholipid cell membranes has been used to form surface patterns. Zhang et al. [58] prepared patterns of a kind of adhesion peptide (RADS) on the gold surface using a combination of a self-assembly oligopeptide monolayer and microcontact printing. Oligopeptides containing a cell adhesion motif (RADS) at the *N* terminus, followed by a hydrophobic oligo–alanine linker and a cysteine residue at the *C* terminus were prepared. By μCP, a patterned SAM of the oligopeptides was formed on the gold surface. The —SH group can react with the gold to attach the oligopeptide monolayer covalently to the gold surface. Hexaethylene glycol thiolate was used to produce antiadhesive domains. Microscale RADS tracks were prepared to guide cell adhesions and orientations precisely (Fig. 18 [58]).

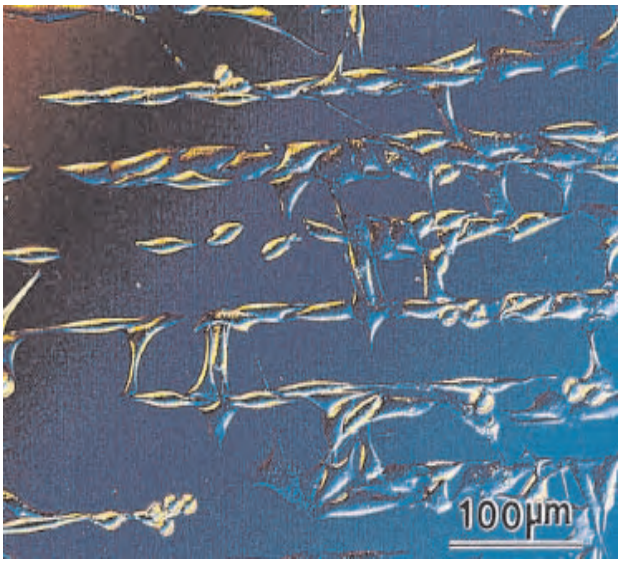


Figure 18. Mouse fibroblast NIH 3T3 cells in linear arrays. Reprinted with permission from [58], S. Zhang et al., *Biomater.* 20, 1213 (1999). © 1999, Elsevier Science.

5. NANOFIBERS AND 3D NANOSCALE SYNTHETIC EXTRACELLULAR MATRIX

The ideas that use biodegradable 3D nonwoven meshworks fabricated with synthetic polymeric nanofibers as tissue-engineering scaffolds have been attracting more and more interest in recent years. This is for three reasons: (1) polymeric fibers are usually strong in tensile strength, so the scaffold can have good mechanical properties and excellent flexibility; (2) research showed that cells attach and organize well around fibers with diameters smaller than the diameter of the cells [59]; and (3) a nonwoven polymeric meshwork mimics the nanoscale protein fiber meshwork in natural ECM.

5.1. Cells and Fibers

Wan et al. [60] studied the fibroblast adhesion and proliferation on four different multifilament materials with diameters less than 20 μm: Kevlar 29 (K), silicon carbide (SiC), nylon 66 (N), and polyethylene terephthalate (PET). The results showed that fibroblasts adhered and grew very well on all fiber surfaces, although fewer cells were observed on PET. Significant orientational behavior of the cells was found on all fiber surfaces from the SEM and CLSM analysis, independent of the bulk chemistry of the fibers. To study the effects of nanofiber diameter on osteoblast proliferation and differentiation, carbon fiber nonwoven mesh with different fiber diameters (from 60 to 200 μm) were synthesized [61]. Osteoblast proliferation increased with decreasing carbon fiber diameters after three and seven days of culture. Moreover, compared to larger diameter carbon fibers, osteoblasts synthesized more alkaline phosphatase and deposited more extracellular calcium on the carbon fibers with smaller diameters.

5.2. Electrospinning

One of the most important technologies to prepare polymeric nanofibers is electrostatic spinning (electrospinning). Electrospinning is not a new technology for polymer fiber production. It has been known since the 1930s [62]. In recent years, special needs of military, medical, and filtration applications have stimulated renewed interests and studies in the process of electrospinning. The principle of electrospinning is to use an electric field to draw a positively charged polymer solution or melt from an orifice to a collector (Fig. 19 [63]). This creates a jet of fluid that is drawn down by acceleration from the orifice to the grounded collection device. The process use high voltages (10–20 kV) to generate sufficient surface charge to overcome the surface tension in a pendant drop of the polymer fluid. The diameters of electrospun fibers are at least one order of magnitude smaller than those made by conventional extrusion techniques. When the jet dries or solidifies, an electrically charged fiber remains. This charged fiber can be directed by electrical forces, and then deposited in sheets or other useful geometrical shapes in the form of a nonwoven fabric. Due to the small fiber diameters and the porous structure of a nonwoven meshwork, electrospun products possess a high surface-to-volume ratio that renders them attractive for a variety of applications, such as a biodegradable nonwoven meshwork for tissue engineering. Figures 20 [64] and 21 [65] are SEM photographs of polymeric nanofibers prepared by the electrospinning method.

Reneker et al. [66] prepared nanofibers of more than 20 kinds of polymers, including polyethylene oxide, nylon, polyimide, DNA, polyaramid, and polyaniline using electrospinning. Most fibers were spun from polymer solution. In another work by Fong and Reneker et al. [67], styrene–butadiene–styrene triblock copolymers were electrospun to obtain elastic nanofibers with diameters around 100 nm. Separated phases of styrene and butadiene blocks were observed using TEM. The single-phase domains were irregular in shape, but elongated along the axis of the fiber. Biodegradable amorphous poly(D,L–lactic acid) (PDLA)

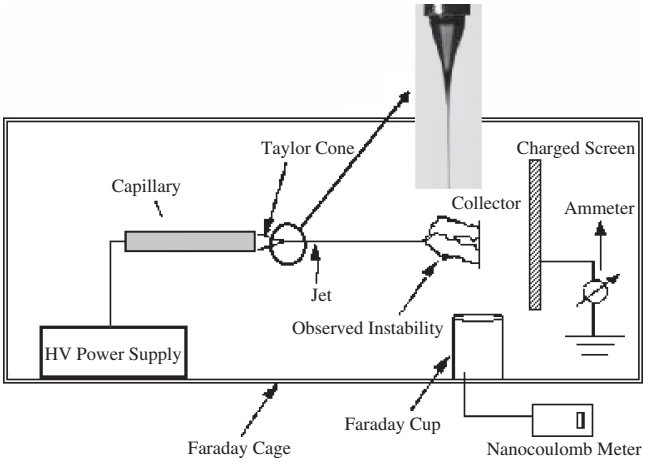


Figure 19. Schematic diagram of electrospinning. Reprinted with permission from [63], G. C. Rutledge et al., *National Textile Center Annual Report* M01-D22, 2001. © 2001, National Textile Center.

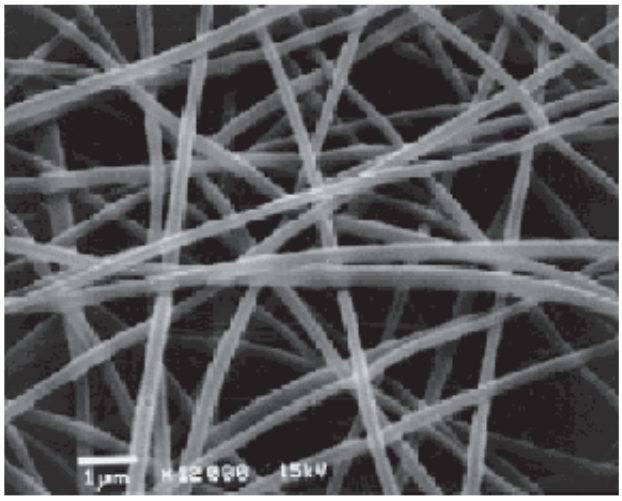


Figure 20. SEM photograph of electrospun polyacrylonitrile fiber. Reprinted with permission from [64], S. B. Warner et al., *National Textile Center Annual Report* M98-D01, 1998. © 1998, National Textile Center.

and semicrystalline poly(L–lactic acid) (PLLA) nanofiber nonwoven membranes for biomedical applications were fabricated by Zong et al. [68] using the electrospinning method. Bognitzki et al. [65] prepared PLLA, PC, and polyvinylcarbazole nanofibers with pores on fiber surfaces. The pore morphology can be controlled by the rapid phase separation induced by the evaporation of the solvent and a subsequent rapid solidification. Not only synthetic, but also natural polymers such as protein, have been made into nanofibers. Huang et al. [22] employed a kind of elastin-mimic protein that was processed into fibers by an electrospinning technique. The 81 kDa recombinant protein based upon the repeating elastomeric peptide sequence of elastin (Val–Pro–Gly–Val–Gly)(4)(Val–Pro–Gly–Lys–Gly) was obtained through bacterial expression of the gene

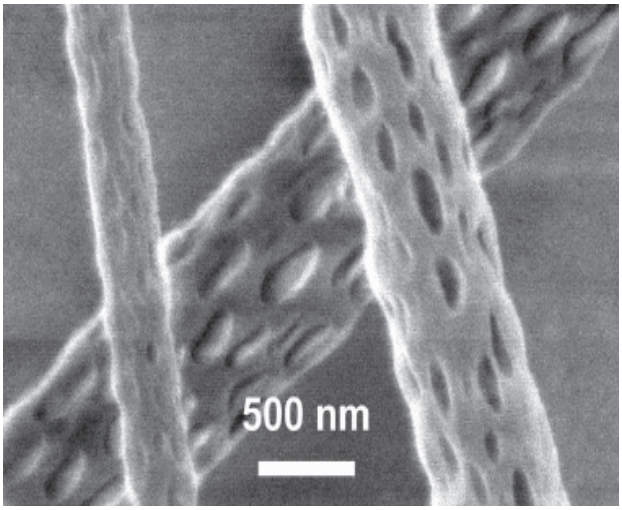


Figure 21. SEM micrographs of porous PLLA fibers obtained via electrospinning of a solution of PLLA in dichloromethane. Reprinted with permission from [65], M. Bognitzki et al., *Adv. Mater.* 13, 70 (2001). © 2001, Wiley-VCH.

coding for tandem repeats of the monomer. Processing parameters were changed to influence the fiber diameter (200–3000 nm) and morphologies. Three kinds of morphologies, beaded fibers, thin filaments, and broad ribbon-like structures, were obtained with varied processing parameters.

Using a PEO/water system, Deitzel et al. [69] studied the effects of the two most important parameters, spinning voltage and solution concentration, on the morphology of the fibers formed. In addition, electrostatic effects also are found to influence the macroscale morphology of the electrospun textiles, and may result in the formation of heterogeneous or three-dimensional structures. Mo et al. [70] found that the diameter of PLLA–co–PCL nanofibers decreased with decreasing polymer concentration and increasing electrospinning voltage. Other affecting parameters include polymer solution vescoelasticity, polymer solution feeding rate, and ionic slat addition [68].

One of the characteristic features of electrospinning is the onset of a chaotic oscillation of the electrospinning jet. This instability of the polymer solution jet can cause the formation of beaded fibers. Fong and Reneker [71] demonstrated that electrospun fibers often have beads in regular arrays (Fig. 22 [71]). The viscoelasticity of the solution, charge density carried by the jet, and the surface tension of the solution are the key factors that influence the formation of the beaded fibers. Reneker et al. [72] studied the bending instability of the electrically charged liquid jets of polymer solutions using a mathematical rheological model. Deitzel et al. [73] overcame the instability by using an electrostatic lens element and a collection target of opposite polarity to control the deposition of PEO fibers (diameters <300 nm) on a substrate.

5.3. Phase Separation

Phase separation method has been used to prepare 3D nanofiber meshworks. Using this technique, Ma et al. [74] prepared biodegradable 3D meshworks composed of PLLA, PDLA, and PLGA nanofibers with diameters ranging from 50 to 500 nm. The process involved thermally induced gelation (phase separation), solvent exchange, and freeze drying. It should be noted that a sponge-like porous matrix could also be obtained by the phase separation method [75]. Preparing parameters such as gelation temperature, gelation time, solvent, and polymer concentration should be well controlled to make proper conditions for the formation of nanofibers in the phase separation to obtain a nanofiber meshwork morphology.

The main advantage of the phase separation method is that it is easy to fabricate meshworks that are isotropic in all three dimensions. It is also a simple process requiring little equipment. However, the fibers produced are relatively short, and do not have a uniform cross-sectional geometry or porosity along their lengths. Hence, the scaffolds lack mechanical strength, and it may be difficult to produce scaffolds with consistent and controlled architecture.

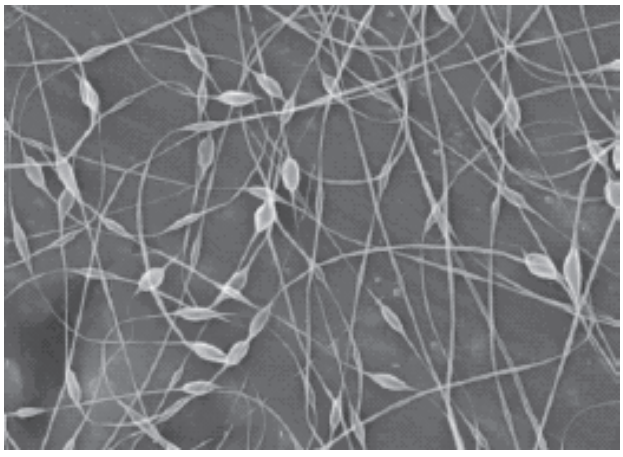


Figure 22. Beaded nanofibers. The horizontal edge of each image is 20 μm long. Reprinted with permission from [71], H. Fong et al., *Polym.* 40, 4585 (1999). © 1999, Elsevier Science.

5.4. Nonwoven Meshworks as Tissue-Engineering Scaffolds

Nonwoven meshes of PLGA or PGA fibers with a diameter of 10–15 μm have been widely used in cartilage tissue engineering [76]. Cao et al. [77] seeded chondrocytes isolated from bovine articular cartilage into PGLA nonwoven meshworks with the shape of a human auricle. The cell-seeded nonwoven meshworks were subsequently implanted on the dorsa of athymic mice. New cartilage tissue was generated in the meshworks. The three-dimensional structure of the meshworks was well maintained in the process of the cartilage tissue regeneration after the removal of the external stents. PGA nonwoven mesh with a diameter of 12–15 μm has also been used for cardiovascular tissue engineering by first seeding the fibroblast and then epithelia cells on the mesh [78]. Microscopic examination of the seeded mesh demonstrated that the human fibroblasts attached to the polymeric fibers, and spread out and proliferate in the meshwork. The seeded endothelial cells formed a monolayer on the fibroblasts (Fig. 23 [78]). Finally, a homogeneous scaffold resembling a solid sheet of tissue was obtained. Li et al. [79] developed a thermal compression method to prepare a nonwoven polyethylene terephthalate (PET) meshwork with a fiber diameter of about 20 μm. The effects of preparation parameters including temperature, pressure, and compression duration on the mechanical property, pore size, and porosity of the meshworks were studied. Human trophoblast ED_{27} and NIH 3T3 cells were cultured and adhered on the PET fibers (Fig. 24 [79]). This thermal compression method can be used to prepare other biodegradable polymer nonwoven mesh, and to modify their microstructure. Altman et al. [80] used a natural silk cord matrix (5–20 μm in diameter) as a scaffold for the tissue engineering of anterior cruciate ligaments. The silk fiber matrix promoted the adhesion, spreading, and differentiation toward the human ligament lineage (Fig. 25 [80]).

Fewer works have been reported to prepare nonwoven meshworks using nanoscale fibers for tissue engineering, probably due to the fact that fibers with such small diameters are very difficult to easily manipulate. Mo et al. [70] found that smooth muscle cells (SMCs) and endothelia cells

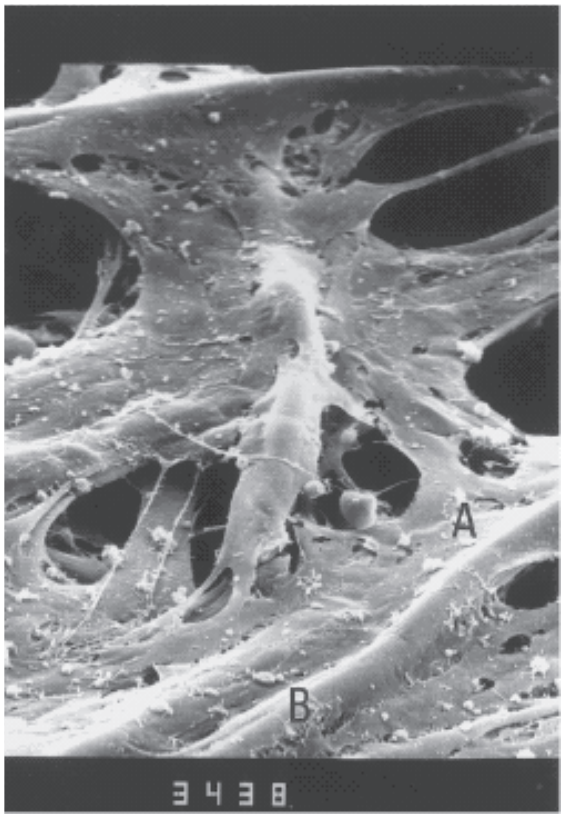


Figure 23. Continuous formation of fibroblasts. Scanning electron microscopy shows a continuous distribution of fibroblasts throughout the "polymeric architecture" and attachment of the fibroblasts (A) to the polymer fibers (B). Reprinted with permission from [78], G. Zund et al., *Eur. J. Cardio-Thoracic Surg.* 13, 160 (1998). © 1998, Elsevier Science.

(ECs) adhered, proliferated, and spread well on PLLA–co–PCL nanofiber meshworks obtained by electrospinning. In a recent work reported by Vacanti et al. [81], a PCL nanofiber nonwoven meshwork obtained by electrospinning

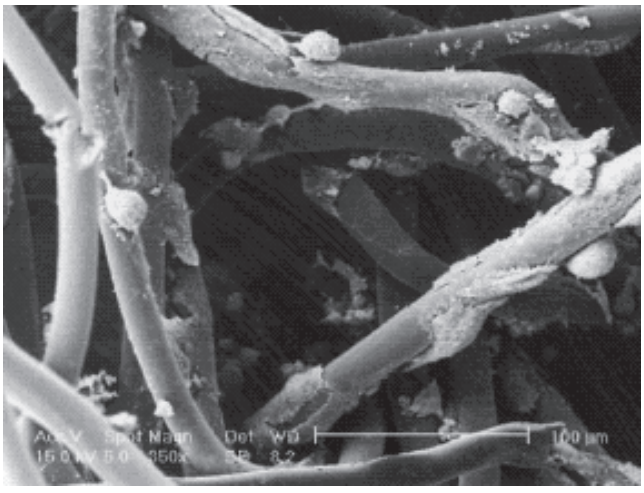


Figure 24. SEM pictures of human trophoblast ED27 cultured in nonwoven PET fabrics. Reprinted with permission from [79], Y. Li et al., *Biomater.* 22, 609 (2001). © 2001, Elsevier Science.

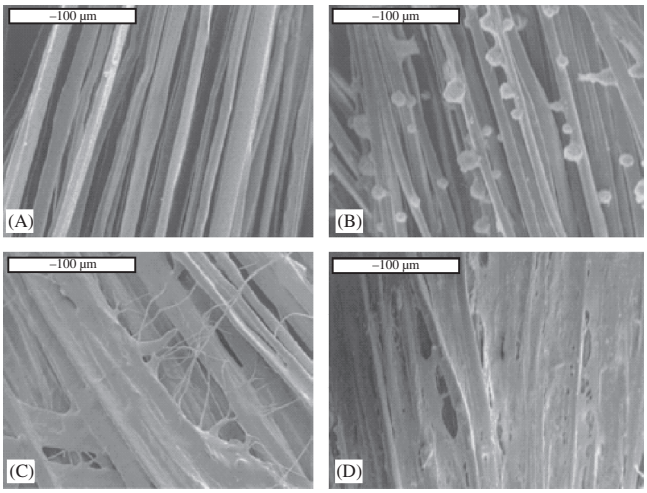


Figure 25. SEM showing adherence, proliferation, and cell sheet formation by human progenitor bone marrow stromal cells on the silk cord matrix prior to seeding (A), time 0 following seeding (B), 1 day (C), and 14 days (D). Scale bars = 100 mm. Reprinted with permission from [80], G. H. Altman et al., *Biomater.* 23, 4131 (2002). © 2002, Elsevier Science.

was studied for its potential in bone tissue engineering. Mesenchymal stem cells (MSCs) derived from the bone marrow of neonatal rats were cultured, expanded, and seeded on the PCL meshwork. The cell–meshwork construct was then cultured in a culture medium supplemented with osteogenic factors under dynamic culture conditions for more than four weeks *in vitro*. The cells penetrated into the meshwork and secreted ECM during the culture. After four weeks of culture, collagen type I and mineralized ECM were found in the cell–meshwork construct, which indicated that the electrospun PCL is a potential candidate scaffold for bone tissue engineering.

It can be expected that, in the future, nonwoven meshworks composed of nanoscale fibers will be more and more widely used as ECM analogs in the field of tissue engineering and artificial organs.

6. CONCLUSION

In natural ECM, protein fibers with nanoscale diameters form a nonwoven meshwork penetrating through and embedded in the proteoglycan hydrogel. In common connective tissue, the fiber diameter is from several hundred nanometers to several microns, while in basal lamina, the fibers are thinner (several to several ten nanometers in width) and the meshwork is denser (several nanometers in pore diameter). The interactions between the cells and the ECM are mainly mediated by the integrin-clustered focal contacts, of which the dimensions are also from several hundred nanometers to several microns. These nanometric characteristics of an ECM can be used to prepare ECM-mimic synthetic materials. Nano/micropatterns can be produced on plane biomaterial surfaces using several kinds of microconstruct technologies such as photolithography, microcontact printing, and self-assembly monolayer. The nano/microsurface patterns have significant influences on

cell behavior. Polymeric nanofibers can be obtained successfully by electrospinning technology, and fabricated into nonwoven meshworks. These nonwoven meshworks have a potential application in tissue engineering in the future due to their great similarity with the nonwoven meshworks in a natural ECM.

GLOSSARY

Adhesion proteins (also called adhesive proteins) Fibrous proteins in ECM which promote cell adhering on ECM by specifically binding cell membrane receptors called integrin. Adhesion proteins mainly include fibronectin, laminin and vitronectin etc. Collagen is also a kind of adhesion proteins due to its ability to bind with some integrins.

Biomaterials Materials which directly contact physiological environments like cells, blood and human tissues.

Electrospinning A polymer processing method to produce ultra-fine polymer fibers. The principle of electrospinning is to use a high electric field to draw a positively charged polymer solution or melt from an orifice to a collector, producing ultra-fine fibers with diameters ranging from several microns down to several tens of nanometers.

Extracellular matrix (ECM) A complex structural entity surrounding and supporting cells that are found within mammalian tissues.

Focal contact (focal adhesion) Focal contact, also called focal adhesion, is formed when a cluster of integrins accumulate in a relatively small spot on the cell membrane and interact with adhesion proteins in ECM. Focal contact is the main way for cells to attach onto the ECM and to respond to outer signals.

Glycosaminoglycans (GAGs) Long unbranched polysaccharides containing a repeating disaccharide unit. The disaccharide units contain either of two modified sugars—*N*-acetylgalactosamine (GalNAc) or *N*-acetylglucosamine (GlcNAc) and a uronic acid such as glucuronate or iduronate. The majority of GAGs in ECM are linked to core proteins, forming proteoglycans.

Integrin Tans-membrane heterodimeric receptor proteins on cell surface. There are about 22 different kinds of integrins on cell surface. The binding of integrins with the adhesive proteins lead to the forming of focal contact, which is the main way for cell to adhere onto the ECM.

Nanofiber Fibers with diameter ranging from several nanometers to several hundred nanometers are called nanofiber in this chapter.

Proteoglycan Macromolecules that are composed of a protein core to which is attached many GAGs chains. Proteoglycan hydrogels provide pressure resistance properties for ECM.

Structural proteins Fibrous proteins in ECM which provide tensile strength and elasticity for ECM. Structural proteins mainly include collagen and elastin. Collagen is the most abundant protein in ECM.

Surface patterning Surface treatment of materials which can produce micro- or nano-scale texture with a random or regular geometric pattern on the material surfaces. The most common surface patterning techniques include photolithography electrobeam-lithography, neutron-lithography, self-assembly and microcontact printing, etc.

Synthetic ECM Artificial materials that can support the growth of cells, just as the native ECM does.

Tissue engineering A new approach to repair and regenerate damaged human tissues. The procedure include seeding the cells into a 3-D porous biodegradable scaffolds, culturing the cell-scaffold construct *in vitro*, and transplanting the cell-scaffold construct into human body. The scaffold will be degraded and removed, and finally replaced by new generated tissue, living no foreign materials *in vivo*.

Tissue engineering scaffold 3-D porous biodegradable biomaterial used as supports for cells in tissue engineering.

REFERENCES

1. L. L. Hench and J. M. Polak, *Science* 295, 1014 (2002).
2. C. H. Lee, A. Singla, and Y. Lee, *Int. J. Pharmaceut.* 221, 1 (2001).
3. N. V. Majeti, Kumar, and Ravi, *Reactive Functional Polym.* 46, 1 (2000).
4. B. Grigolo, L. Roseti, M. Fiorini, M. Fini, G. Giavaresi, N. N. Aldini, R. Giardino, and A. Facchini, *Biomater.* 22, 2417 (2001).
5. Loty, S. Sautier, J. M. Loty, C. Boulekbache, H. Kokubo, and T. Forest, *J. Biomed. Mater. Res.* 42, 213 (1998).
6. S. Kouvroukoglou, K. C. Dee, R. Bizios, L. V. McIntire, and K. Zygourakis, *Biomater.* 21, 1725 (2000).
7. B. L. Seal, T. C. Otero, and A. Panitih, *Mater. Sci. Eng.* R34, 147 (2001).
8. J. Emsley, G. C. Knight, R. W. Farndale, M. J. Barnes, and R. C. Liddington, *Cell* 101, 47 (2000).
9. V. C. Mow, G. A. Ateshian, W. M. Lai, and W. Y. Gu, *Int. J. Solids Struct.* 35, 4945 (1998).
10. B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson, "Molecular Biology of the Cell, Part IV. Cells in Their Social Context," 3rd Ed. Garland, New York, 1994.
11. http://www.rcsb.org/pdb/molecules/pdb4_3.html.
12. http://www.virtuallaboratory.net/BioFun-KFI/lectureNotes/Topic6-3_Adhesion.htm.
13. S. Mosler, D. Selmeczi, N. B. Larsen, N. Gadegaard, P. Kingshott, S. M. Arnesen, and M. A. Lawson, "Symposium on Nano-Physics of Life Systems," Copenhagen, June 2002, poster.
14. E. B. Hunziker, W. Herrmann, R. K. Schenk, M. Mueller, and H. Moor, *J. Cell Biol.* 98, 277 (1984).
15. T. Nishida, K. Yasumoto, T. Otori, and J. Desaki, *Invest. Ophthalmol. Vis. Sci.* 29, 1887 (1998).
16. R. G. Flemming, C. J. Murphy, G. A. Abrams, S. L. Goodman, and P. F. Nealey, *Biomater.* 20, 573 (1999).
17. I. Shirato, Y. Tomino, H. Koide, and T. Sakai, *Cell Tissue Res.* 266, 1 (1991).
18. Y. Yamasaki, H. Makino, and Z. Ota, *Nephron* 66, 189 (1994).
19. K. Hironaka, H. Makino, Y. Yamsaki, and Z. Ota, *Kidney Int.* 43, 334 (1993).
20. G. A. Abrams, S. L. Goodman, and P. F. Nealey, *Invest. Ophthalmol. Vis. Sci.* 38, 350 (1997).
21. G. A. Abrams, S. L. Goodman, P. F. Nealey, and C. J. Murphy, *Invest. Ophthalmol. Vis. Sci.* 39, 5160 (1998).
22. L. Huang, B. A. McMillan, R. P. Apkarian, B. Pourdeyhimi, V. P. Conticello, and E. L. Chaikof, *Macromol.* 33, 2989 (2000).
23. http://ntri.tamuk.edu/homepage-ntri/lectures/protein/junction1.html.
24. M. Tirrell, E. Kokkoli, and M. Biesalski, *Surf. Sci.* 500, 61 (2002).
25. http://www.virtuallaboratory.net/BioFun-KFI/lectureNotes/Topic6-3_Adhesion.htm.

26. K. Burridge, K. Fath, T. Kelly, G. Nuckolls, and C. Turner, *Ann. Rev. Cell Biol.* 4, 487 (1988).
27. J. S. Burmeister, L. A. Olivier, W. M. Reichert, and G. A. Truskey, *Biomater.* 19, 307 (1998).
28. C. S. Chen, M. Mrksich, S. Huang, G. M. Whitesides, and D. E. Ingber, *Science* 276, 1425 (1997).
29. J. M. Schakenraad, H. J. Busscher, C. R. H. Wildevuur, and J. Arends, *J. Biomed. Mater. Res.* 20, 773 (1986).
30. N. J. Hallab, K. J. Bundy, K. O'Connor, R. L. Moses, and J. J. Jacobs, *Tissue Eng.* 7, 55 (2001).
31. P. B. V. Wachem, T. Beugelling, J. Feijen, A. Bantjes, J. P. Detmers, and W. G. V. Aken, *Biomater.* 6, 403 (1985).
32. Z. Ma, C. Gao, and J. Shen, *J. Biomater. Sci. Polym. Ed.* 14, 13 (2003).
33. P. B. V. Wachem, A. H. Hoget, T. Beugeling, J. Feijen, A. Bantjes, J. P. Detmers, and W. G. V. Aken, *Biomater.* 8, 323 (1987).
34. J. H. Lee, J. W. Lee, G. Khang, and H. B. Lee, *J. Biomed. Mater. Res.* 18, 351 (1997).
35. J. H. Lee, G. Khang, and H. B. Lee, *J. Biomed. Mater. Res.* 40, 180 (1998).
36. J. H. Lee, H. W. Jung, I. K. Kang, and H. B. Lee, *Biomater.* 15, 705 (1994).
37. Z. Ma, C. Gao, Y. Gong, and J. Shen, *Biomater.* (2003).
38. M. Lampin, R. Warocquier, C. Legris, M. Degrange, and M. F. Sigot-Luizard, *J. Biomed. Mater. Res.* 36, 99 (1997).
39. D. D. Deligianni, N. D. Katsala, P. G. Koutsoukos, and Y. F. Missirlis, *Biomater.* 22, 87 (2001).
40. A. Curtis and C. Wilkinson, *Biomater.* 18, 1573 (1997).
41. Y. W. Fan, F. Z. Cui, S. P. Hou, Q. Y. Xu, L. N. Chen, and I. S. Lee, *J. Neurosci. Meth.* 120, 17 (2002).
42. W. R. Wilkerson, C. A. Seegert, A. W. Feinberg, L. C. Zhao, J. A. Callow, M. E. Callow, and A. B. Brennan, *Polym. Preprints* 42, 147 (2001).
43. K. Matsuzaka, X. F. Walboomers, J. E. de Ruijter, and J. A. Jansen, *Biomater.* 20, 1293 (1999).
44. B. Wojciak-Stothard, A. Curtis, W. Monaghan, K. MaCdonald, and C. Wilkinson, *Exp. Cell Res.* 223, 426 (1996).
45. B. Wojciak-Stothard, Z. Madeja, W. Korohada, A. Curits, and C. Wilkinson, *Cell Biol. Int.* 19, 485 (1999).
46. J. Meyle, K. Gultig, M. Brich, H. Hammerle, and W. Nisch, *J. Biomed. Mater. Res.* 29, 81 (1995).
47. A. F. V. Recum and T. G. Van Kooten, *J. Biomater. Sci. Polym. Ed.* 7, 181 (1995).
48. P. Clark, P. Connolly, A. Curtis, J. A. Dow, and C. D. Wilkinson, *Develop.* 99, 439 (1987).
49. P. Clark, P. Connolly, A. S. Curtis, J. A. Dow, and C. D. Wilkinson, *Develop.* 108, 635 (1990).
50. P. Clark, P. Connolly, A. S. Curtis, J. A. Dow, and C. D. Wilkinson, *J. Cell Sci.* 99, 73 (1991).
51. P. Clark, P. Connolly, and G. R. Moores, *J. Cell Sci.* 103, 287 (1992).
52. B. G. Casey, W. Monaghan, and C. D. W. Wilkinson, *Microelectron. Eng.* 35, 393 (1997).
53. A. Mata, C. Boehm, A. J. Fleischman, G. Muschler, and R. Shuvo, *J. Biomed. Mater. Res.* 62, 499 (2002).
54. J. Tan, H. Shen, K. L. Carter, and W. M. Saltzman, *J. Biomed. Mater. Res.* 51, 694 (2000).
55. L. Kam, W. Shain, J. N. Turner, and R. Bizios, *Biomater.* 22, 1049 (2001).
56. R. Singhvi, A. Kumar, G. P. Lopez, G. N. Stephanopoulos, D. I. C. Wang, G. M. Whitesides, and D. E. Inberg, *Science* 264, 696 (1994).
57. G. M. Whitesides and B. Grzybowski, *Science* 295, 2418 (2002).
58. S. Zhang, L. Yan, M. Altman, M. LaKssle, H. Nugent, F. Frankel, D. A. Lauffenburger, G. M. Whitesides, and A. Rich, *Biomater.* 20, 1213 (1999).
59. C. T. Laurencin, A. M. A. Ambrosio, M. D. Borden, and J. A. Cooper, *Annu. Rev. Biomed. Eng.* 1, 19 (1999).
60. H. Wan, R. L. Williams, P. J. Doherty, and D. F. Williams, *J. Mater. Sci.: Mater. Med.* 8, 45 (1997).
61. K. L. Elias, R. L. Price, and T. J. Webster, *Biomater.* 23, 3279 (2002).
62. A. Formhals, U.S. Patent 1, 975, 504, 1934.
63. G. C. Rutledge, Y. Li, S. Fridrikh, S. B. Warner, V. E. Kalayci, and P. Patra, National Textile Center Annual Report M01-D22, 2001.
64. S. B. Warner, A. Buer, S. C. Ugbolue, G. C. Rutledge, and M. Y. Shin, National Textile Center Annual Report M98-D01, 1998.
65. M. Bognitzki, W. Czado, T. Frese, A. Schaper, M. Hellwig, M. Steinhart, A. Greiner, and J. H. Wendorff, *Adv. Mater.* 13, 70 (2001).
66. D. H. Reneker and I. Chun, *Nanotechnol.* 73, 216 (1996).
67. H. Fong and D. H. Reneker, *J. Polym. Sci. Part B: Polym. Phys.* 37, 3488 (1999).
68. X. Zong, K. Kim, D. Fang, S. Ran, B. S. Hsiao, and B. Chu, *Polym.* 43, 4403 (2002).
69. J. M. Deitzel, J. Kleinmeyer, D. Harris, and N. C. B. Tan, *Polym.* 42, 261 (2001).
70. X. M. Mo, C. Y. Xu, M. Kotaki, and S. Ramakrishna, *Biomater.* (submitted, 2003).
71. H. Fong, I. Chun, and D. H. Reneker, *Polym.* 40, 4585 (1999).
72. D. H. Reneker, A. L. Yarin, and H. Fong, *J. Appl. Phys.* 87, 4531 (2000).
73. J. M. Deitzel, J. D. Kleinmeyer, J. K. Hirvonen, and N. C. B. Tan, *Polym.* 42, 8163 (2001).
74. P. X. Ma and R. Zhang, *J. Biomed. Mater. Res.* 46, 60 (1999).
75. Y. S. Nam and T. G. Park, *Biomater.* 20, 1783 (1999).
76. W. S. Kim, J. P. Vacanti, L. Cima, D. Mooney, J. Upton, W. C. Puelacher, and C. A. Vacanti, *Plastic Reconst. Surg.* 94, 233 (1994).
77. Y. Cao, J. P. Vacanti, K. T. Paige, J. Upton, and C. A. Vacanti, *Plastic Reconst. Surg.* 100, 297 (1997).
78. G. Zund, S. P. Hoerstrup, A. Schoeberlein, M. Lachat, G. Uhlschmid, P. R. Vogt, and M. Turina, *Eur. J. Cardio-Thoracic Surg.* 13, 160 (1998).
79. Y. Li, T. Ma, S. T. Yang, and D. A. Kniss, *Biomater.* 22, 609 (2001).
80. G. H. Altman, R. L. Horan, H. H. Lu, J. Moreau, I. Martin, J. C. Richmond, and D. L. Kaplan, *Biomater.* 23, 4131 (2002).
81. H. Yoshimotoa, Y. M. Shina, H. Teraia, and J. P. Vacanti, *Biomater.* 24, 2077 (2003).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanostructured Hybrids from Layered Double Hydroxides

Toshiyuki Hibino

National Institute of Advanced Industrial Science and Technology, Tsukuba, Japan

CONTENTS

1. INTRODUCTION

Layered double hydroxides (LDHs) are a class of lamellar solids, other examples of which include smectite clays, zirconium phosphates, layered titanium dioxides, transition-metal chalcogenides, layered perovskites, graphites, and graphite oxides. LDHs are made up of hydroxide layers with anions and water sandwiched between them. The hydroxide layers vary in composition, and in composition ratio. The weakly interacting regions between the layers are called the "interlayer" or "gallery" regions. LDHs, which are rare in nature but relatively simple and inexpensive to synthesize, are sometimes called "anionic clays," owing to their structural similarity to smectite clays and their anion exchange capability. LDHs go by numerous other names—such as hydrotalcite, manasseite, pyroaurite, sjögrenite, takovite, meixnerite, reevesite, iowaite, and honessite—depending on the composition and polytype form of the minerals. The term "green rust" is used for some natural LDHs found in anoxic environments [1–3]. LDHs are also widely known as "hydrotalcite-like compounds."

LDHs have found a variety of practical applications in industry. Calcination of LDHs at temperature depending on the particular cations in the layers and anions in the interlayers leads to mixed oxides characterized by structural homogeneity and the absence of chemical segregation. Catalysts obtained in this way have been thoroughly investigated, and many papers and reviews have been published [4–6]. LDHs are also being used as antacid agents for gastritis, as stabilizers in poly(vinyl chloride), and as components of flame retardants for plastics. However, all these uses of LDHs are related to either their thermal properties or their basicity and do not take advantage of their two-dimensional lamellar structure.

The technique of delaminating or exfoliating lamellar compounds into single lamellae or packets of a few lamellae is the key route for assembling nanostructured hybrids. Designing and building new materials on a nanometer scale are currently topics of intense research. Single layers of lamellar compounds can be considered as anisotropic inorganic macromolecules, and can serve as well-defined building blocks for the synthesis of heterostructured nanohybrids. For example, smectite clays have found practical use in polymer/clay nanocomposite materials in which the clays have been delaminated into single sheets within polymer matrices [7]. Furthermore, delamination has been reported for several other classes of lamellar compounds, such as zirconium phosphates, layered titanium dioxides, transition-metal chalcogenides, and graphite oxides. These compounds have been thoroughly investigated and have been developed into nanocomposite materials that have already attained success in the experimental stage [8–23]. On the other hand, most of the nanocomposite materials successfully fabricated from LDHs have been obtained by means of techniques other than delamination, as will be discussed later in this review. However, the first example of delamination of an LDH was reported in 2000 [24].

In addition to describing the synthesis and fundamental properties of LDHs, this review compares nanocomposite materials made from LDHs with nanocomposite materials made from other lamellar compounds, focusing on the delamination process.

2. LAYERED DOUBLE HYDROXIDES

2.1. Structure

As we have seen, LDHs consist of positively charged hydroxide layers with interlayer spaces filled with anions and water molecules (Fig. 1) [25, 26]. LDHs can be usually represented

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (657–667)

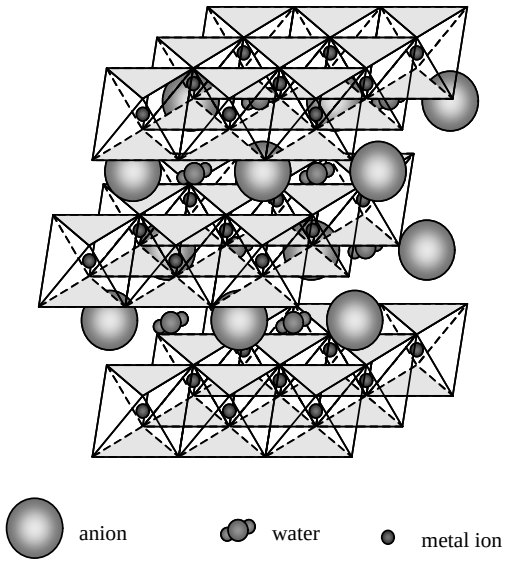


Figure 1. Structure of LDHs.

by the general formula $[M^{2+}_{1-x}M^{3+}_{x}(OH)_2]$ $[A^{n-}_{x/n} \cdot mH_2O]$, where x normally varies between 0.20 and 0.33, M^{2+} and M^{3+} are di$^-$ and trivalent metal ions occupying octahedral positions within the hydroxide layers, A^{n-} is an interlayer anion that compensates for the charge on the layers, and m is the number of interlayer water molecules. Each M^{2+} ion is octahedrally surrounded by six OH^- ions, and adjacent octahedra share edges to form infinite sheets that resemble $Mg(OH)_2$ (brucite) sheets. Isomorphous replacement of some of the M^{2+} ions with M^{3+} ions creates the positive charge on the hydroxide layers. The layers can be stacked in either rhombohedral or hexagonal symmetry. In the rhombohedral stacking sequence, the c parameter of the unit cell is $3c'$, where c' is the combined thickness of one hydroxide layer and one interlayer. In the hexagonal stacking sequence, the c parameter is $2c'$. Synthetic LDHs are normally rhombohedral, and some minerals in nature are hexagonal [4, 5].

In addition to the LDHs represented by the general formula given above, Li–Al LDHs, $[Li^{+}_{1/3}Al^{3+}_{2/3}(OH)_2][A^{n-}_{1/3n} \cdot mH_2O]$, have also been reported [27]. In this case, the metal ions in the host layers are mono- and trivalent cations. Li^+ ions occupy the vacant octahedral sites within the $Al(OH)_3$ (gibbsite) layers in this compound. When Li–Al LDHs are included, the extended general formula for LDHs becomes $[M^{z+}_{1-x}M^{3+}_{x}(OH)_2]^{b+}$ $[A^{n-}_{b/n} \cdot mH_2O]$, where M is a metal ion, A is an anion, and $b = x$ or $2x - 1$ for $z = 2$ or 1, respectively. Intriguingly, although Li–Al LDHs are synthetic, they have hexagonal symmetry, unlike M^{2+}–M^{3+} LDHs.

The phenomenon in which the interlayers of lamellar compounds are partially filled with guest ions or molecules in an ordered way is known as staging. Many examples of staging are found for graphite. Although staging are not commonly found for LDHs, some LDHs have been found to show second-stage intermediate phases when some organic anions are intercalated between the layers [28, 29].

2.2. Synthesis

In general, LDHs are synthesized by coprecipitation [4, 5, 30, 31]. For example, one can obtain a typical Mg-Al-CO_3 LDH by adding an aqueous solution of magnesium and aluminum salts (chloride, nitrate, sulfate, etc.) dropwise to an aqueous solution of sodium carbonate at a constant pH of 10 while simultaneously adding a sodium hydroxide solution [32]. The pH range at which coprecipitation occurs depends on the nature and concentration of the metal ions used (Table 1). The precipitation pH value for the metal ion is calculated from the solubility product (K_{sp}) of $M^{2+}(OH)_2$ or $M^{3+}(OH)_3$ and can serve as a reference for the choice of pH for synthesis. Usually, M^{2+} and M^{3+} double hydroxides coprecipitate in pH titration at pH values between the values at which pure M^{2+} and M^{3+} hydroxides precipitate [4, 33, 34]. The coprecipitation method can be performed at ambient temperature and pressure, and these environmentally benign conditions are one of the virtues of LDHs.

Another virtue of LDH synthesis is the fact that a large variety of compositions are possible, depending on the choice of metal ions and anions. In the coprecipitation synthesis, pure LDH phases are usually obtained at x values ranging from 0.20 to 0.33, and the M^{2+}/M^{3+} atomic ratios in the products are almost the same as those in the initial solutions. There are practically no limitations on the nature of the anions. However, contamination by carbonate anions is likely to occur because LDHs have a high affinity for carbonate anions and because the basic solutions used for LDH preparations easily absorb carbon dioxide from the air [4, 5, 31, 35–37]. When LDHs other than carbonate forms are prepared, measures to avoid carbonate anion contamination should be taken; for example, the synthesis can be carried out under an inert gas like nitrogen or in a glovebox.

2.3. Physical and Chemical Properties

2.3.1. Thermal Behavior

When LDHs are heated, the evolution of interlayer water occurs first and is followed by decomposition to mixed oxides. In the case of the well-studied Mg–Al–CO_3 LDHs

Table 1. Precipitation pH values for metal hydroxides.

		Precipitation pH	
Cation	pK_{sp}	at $[M^{2+\text{ or }3+}] = 0.1$ mol/L	at $[M^{2+\text{ or }3+}] = 0.01$ mol/L
Al^{3+}	32.90	3.4	3.7
Cr^{3+}	30.20	4.3	4.6
Fe^{3+}	37.40	1.9	2.2
Mg^{2+}	10.74	9.1	9.6
Zn^{2+}	16.92	6.0	6.5
Ni^{2+}	14.70	7.2	7.7
Fe^{2+}	15.10	7.0	7.5
Co^{2+}	14.80	7.1	7.6
Cu^{2+}	19.66	4.7	5.2
Ca^{2+}	5.26	11.9	12.4
Mn^{2+}	12.72	8.1	8.6

[4, 32, 33, 38–45], the loss of interlayer water occurs at 100–250 °C and is observed as an endothermic peak in the differential thermal analysis (DTA) curves in Figure 2. This initial loss of water is not accompanied by collapse of the layered structure, and rehydration of the interlayer water occurs reversibly. The second endothermic peak, at 300–400 °C, corresponds to the loss of structural water from the hydroxide layers, as well as decarbonation. At this stage, the layered structure collapses to an amorphous phase, as evidenced by the difference between the X-ray diffraction (XRD) patterns shown in Figure 3a and b. The second endothermic peak in the DTA curves splits into two peaks when the aluminum content increases (Fig. 2). This split is due to the fact that the dehydroxylation of the hydroxyl group bound to aluminum occurs at a lower temperature than the dehydroxylation of the hydroxyl group bound to magnesium. At temperatures higher than 400 °C, a MgO phase begins to appear, and aluminum ions enter the MgO lattice. The calcination materials obtained between 500 °C and 800 °C are NaCl-type Mg–Al oxide solid solutions, as shown in Figure 3c. When the temperature is elevated still further, aluminum ions migrate from the solid solutions to spinel ($MgAl_2O_4$) phase, and at temperatures higher than 900 °C, pure MgO and spinel appear in the XRD pattern (Fig. 3d).

2.3.2. The Memory Effect

When oxide solid solutions obtained by thermal decomposition of Mg–Al–CO_3 LDHs at about 500 °C are exposed to water, the original LDH structure is restored. This reconstruction process, depicted schematically in Figure 4, is termed the "memory effect" [4]. LDHs with volatile interlayer anions, such as chloride and hydroxide anions, also exhibit the memory effect. Because the oxide solid solutions adsorb anions as well as water from the aqueous solution, a great variety of anions can be intercalated by means of the reconstruction method, and numerous LDH intercalates have been reported [4, 30, 31, 46–49]. However, the spinel phase that forms at temperatures higher than 900 °C cannot be reconstructed to LDHs by means of this method.

The reconstruction reaction is mainly governed by the nature of the divalent metal ions in the hydroxide layers of the LDHs. Oxide solid solutions obtained from LDHs based on Mg or Zn hydroxides are easy to reconstruct at room

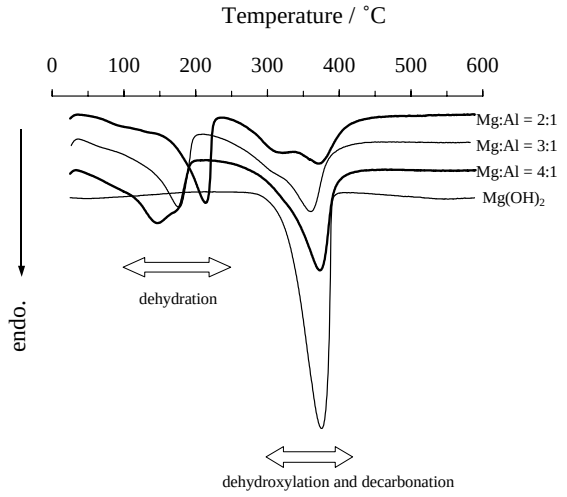


Figure 2. DTA curves of Mg–Al–CO_3 LDHs with various Mg:Al atomic ratios.

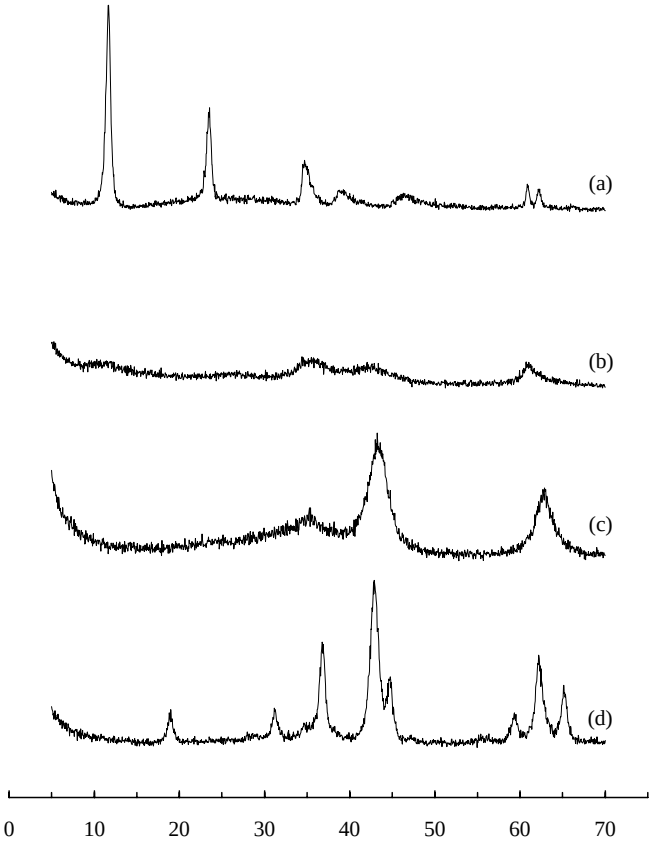


Figure 3. XRD patterns of Mg–Al–CO_3 LDHs (Mg:Al = 2:1) heated at various temperatures: (a) unheated, (b) 350 °C, (c) 500 °C, and (d) 1000 °C.

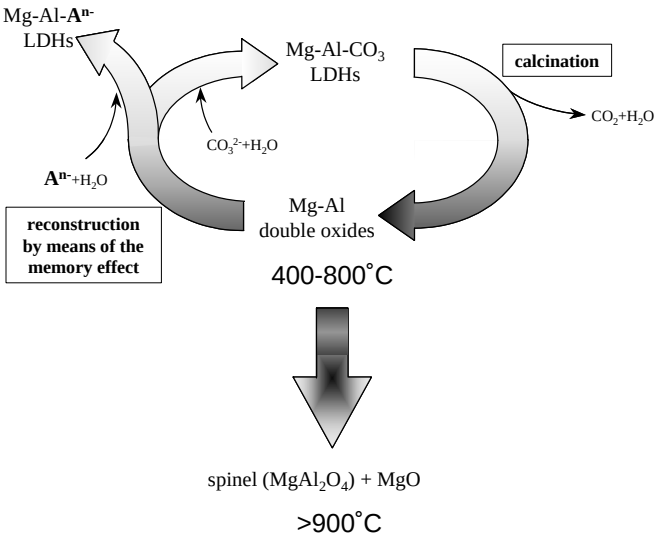


Figure 4. Thermal decomposition and reconstruction of LDHs.

temperature [39, 46–52]. However, reconstruction of oxide solid solutions obtained from LDHs based on Ni hydroxide requires a hydrothermal treatment [52, 53]. Some extent of ZnO-like phase remains after reconstruction at room temperature in the case of Zn-based LDHs, but hydrothermal treatment for Zn-based oxide solid solutions predominantly produces highly crystalline ZnO, and reconstruction into LDHs does not occur [50].

2.3.3. Anion Exchange

One of the most important properties of LDHs is their high anion exchange capacity, which is related to their lamellar structure. When LDHs are dispersed in aqueous solutions containing anions, this capacity allows the exchange of the LDHs' original anions with those present in the solution. This is, of course, one route for making nanostructured hybrids. Anion exchange capacity values ranging from 1.8 to 4.5 mequiv/g have been reported [5, 36, 54, 55], depending on the x value in the general formula. The selectivity of the exchange decreases with decreasing anion charge density. Therefore, LDHs containing monovalent anions, such as chloride and nitrate anions, should be good precursors for ion exchange reactions.

Although carbonate forms of LDHs are easier to synthesize than other anion forms, carbonate anions in LDHs cannot be exchanged with other anions by simple dispersion of a CO_3-LDH in an aqueous solution, because of the high affinity of carbonate for LDHs. Treatment with a diluted acid solution, however, makes the exchange of carbonate anions with other anions possible [37, 56–58]. Chloride and nitrate anions are intercalated when hydrochloric and nitric acids, respectively, are used.

For confirming the presence of anions or polymeric anions in LDHs and determining their environment, IR spectroscopy is a useful technique. For example of Mg–Al–CO_3 LDH, an intense absorption band for carbonate anions appears at 1350-1380 cm^{-1}. After loss of interlayer water, the band splits into two bands, at 1330 and 1540 cm^{-1}, owing to the change of symmetry of the carbonate anions (see Fig. 5).

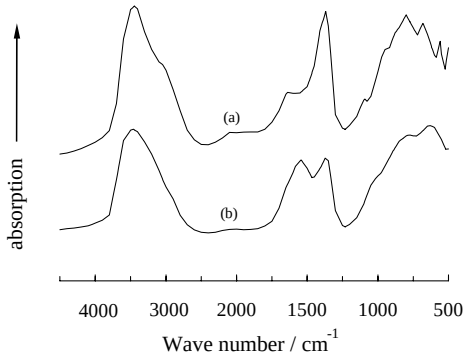


Figure 5. IR spectra of Mg–Al–CO_3 LDHs (Mg:Al = 2:1): (a) unheated and (b) dehydrated at 250 °C. The ν_3 bands of carbonate are labeled with solid circles.

3. NANOSTRUCTURED HYBRIDS FROM LAYERED DOUBLE HYDROXIDES

3.1. Delamination Techniques

Delamination, a process that is unique to lamellar compounds, is the most useful technique for building nanocomposite materials. As we have seen, lamellar compounds allow intercalation of ions or polar molecules into interlayer galleries with the layered structure intact, owing to the considerable anisotropy in the chemical bonding. Large guests open up the interlayer galleries, but the lamellar solids remain crystalline because of the attractive interactions between the guests and the sheets of the host. In some cases, the intercalation process allows large amounts of solvent to penetrate the interlayer region, in a process known as swelling. Delamination or exfoliation of lamellar solids into single sheets can be thought of as the extreme case of swelling, in other words, infinite swelling, as illustrated in Figure 6. The delaminated single sheets can be seen as giant planar macromolecules and as building blocks for new materials that cannot be prepared by means of other processes. This feature of delamination has prompted use of the lamellar solids as precursors for polymer/solid sheet nanocomposite materials [8–16], self-assembled multilayers [17, 18], and pillared porous materials [19–23].

The delamination of lamellar solids other than LDHs has been extensively investigated. Delamination has been reported for smectite clays [14, 59–61]; graphite oxides [9, 10, 62]; dichalcogenides (MoS_2 [11, 63–65], WS_2 [64], and $NbSe_2$ [8]); metal phosphates (α-metal phosphates [18, 66–74], γ-zirconium phosphates [20, 21, 75], and vanadium phosphorus oxides [76]); and layered oxides (layered titanate [16, 17, 22, 77–84], perovskites [85–88], MoO_3 [12, 13], and $K_4Nb_6O_{17}$ [89]). Smectite clays are a typical example, and delamination occurs when suspensions of their sodium

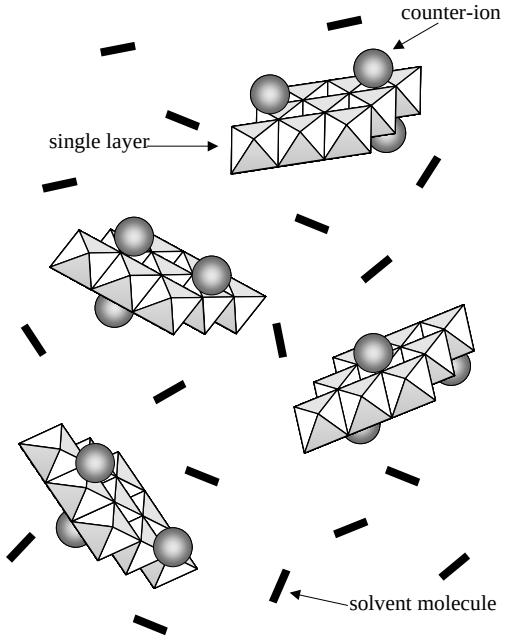


Figure 6. Delamination of LDHs in solvent.

forms are stirred in water. The exfoliation of graphite oxides easily occurs in dilute alkaline aqueous solution. The lithium forms of dichalcogenides have been shown to exfoliate into single sheets upon reaction with water. As for metal phosphates, α-M^{4+} phosphates (M^{4+} = Zr, Sn, and Ti) are known to delaminate in aqueous solutions of alkylamines or tetra-*n*-butylammonium hydroxide (TBA^+OH^-). Layers of γ-zirconium phosphates have been found to delaminate in aqueous alkylamine solutions and in a 50% (v/v) water/acetone mixture at 80 °C. Vanadium phosphorus oxides, $VOPO_4 \cdot 2H_2O$ and $VOHPO_4 \cdot 0.5H_2O$, are reportedly intercalated with 4-butylaniline, and the intercalates are delaminated in tetrahydrofuran (THF). In addition, some layered oxides can be also delaminated. Perovskite oxides such as Dion-Jacobson and Ruddlesden-Popper phases have been successfully delaminated in aqueous solutions of organic bases (amines or amino acids), and delaminations of layered titanates and acid-exchanged $K_4Nb_6O_{17}$ by reaction with aqueous TBA^+OH^- solutions have been confirmed. Some Aurivillius-type perovskite oxides and MoO_3 exfoliate in water after being converted to their lithium forms.

Despite the importance of delamination, techniques for characterizing the process have not been established yet, owing to the difficulty of measuring nanometer-scale phenomena in solutions. XRD measurement of the colloidal solutions in which the lamellar solids are delaminated can sometimes give useful information, although the small and weak peaks of the lamellar compounds are difficult to distinguish from the large and broad peaks from the media. In spite of these difficulties, the XRD pattern of single layers of MoS_2 in water [65] and the XRD patterns of intermediate states such as osmotically swollen layered titanates [82, 84] have been reported. Osmotic swelling had been observed only for smectite clays before being found in layered titanates [59–61, 90]. In the state of osmotic swelling, uptake of much larger volumes of water takes place than does in crystalline swelling.

In addition to XRD, transmission electron microscopy (TEM) and atomic force microscopy (AFM) are other possible characterization methods for delamination. Single sheets obtained by dipping TEM grids or substrates into colloidal solutions in which the lamellar solids are delaminated have been observed by TEM and AFM, and their thickness has been measured by AFM [18, 66, 78, 88]. In the case of tin phosphates, the absence of free phosphoric acid in the colloidal solutions obtained by delamination has been confirmed by ^{31}P NMR spectroscopy, which suggests exfoliation rather than chemical decomposition [91].

Investigation of the delamination of LDHs began only recently. Although the high charge density on LDH layers makes delamination difficult, LDHs have been delaminated by modification of interlayer galleries with organic anions. Delamination of dodecyl sulfate intercalates of Zn–Al LDHs has been reported to occur in refluxing alcohols at 120 °C [24]. A method whereby LDHs containing glycine are delaminated in formamide at room temperature has also been reported [92]. Intercalation of glycine, the simplest amino acid, creates an interlayer region that is attractive to polar solvents, owing to the strong hydrogen bonding between the intercalated anions and the solvents. After publication of these two studies ([24] and [92]), delamination of LDHs has been achieved by means of the shearing of dodecyl sulfate-LDHs in an acrylate monomer at 70 °C and subsequent polymerization of the monomer [93]. Delamination by means of these techniques is becoming a route for the construction of nanocomposite materials from LDHs.

3.2. Polymer/Layered Double Hydroxide Nanocomposite Materials

3.2.1. Synthesis

Nanocomposites composed of polymers and lamellar compounds have been extensively studied with the goal of finding novel functionalities that result from synergistic effects between the polymers and hosts. Novel anisotropic materials, battery cathode materials, electrochromic displays, microelectronic devices and sensors, mechanically reinforced polymers, and systems for the study of polymerization processes have been reported [8–14, 94, 95]. Synthetic methodologies vary with the combination of guest and host but can be placed in four principal categories [96] (see Fig. 7):

(1) *in-situ* polymerization methods
(2) template methods
(3) delamination–reaggregation methods
(4) melt intercalation methods.

Although all four methods should be applicable to LDHs, the first two have been the main processes for LDHs up to now. The *in-situ* polymerization method is feasible only when appropriate monomers are available. In the case of LDHs, monomer anions can be intercalated by ion exchange. Intercalated monomers can then be polymerized by a thermal, photoinduced, or redox treatment. There are two possibilities for the timing of polymerization: polymerization after intercalation of the monomers, and concomitant polymerization during intercalation of the monomers. The template method involves the synthesis of lamellar compounds in solutions in the presence of polymers. During the nucleation and growth of the lamellar host crystals, polymers get trapped within the interlayer regions. LDHs are remarkably suitable for this technique because they can be synthesized at room temperature, whereas other lamellar compounds must be prepared at high temperatures, at which polymers would decompose. In the delamination–reaggregation method, lamellar compounds are delaminated in solvent, and then solutions of polymers are added. Nanocomposites are obtained after evaporation of the solvents, or when resulting nanocomposites are insoluble, they are obtained as precipitates during the reactions. In the melt intercalation method, which has been widely used for smectite clays, lamellar compounds are mixed with melted polymers, and polymer chains diffuse into the interlayer regions.

3.2.2. Nanocomposites

Conventional composites have been prepared with a variety of inorganic additives as stabilizers, extenders, or reinforcements for improving the stiffness and strength of polymers. However, these additives sometimes have drawbacks, such as brittleness or opacity. In these conventional composites,

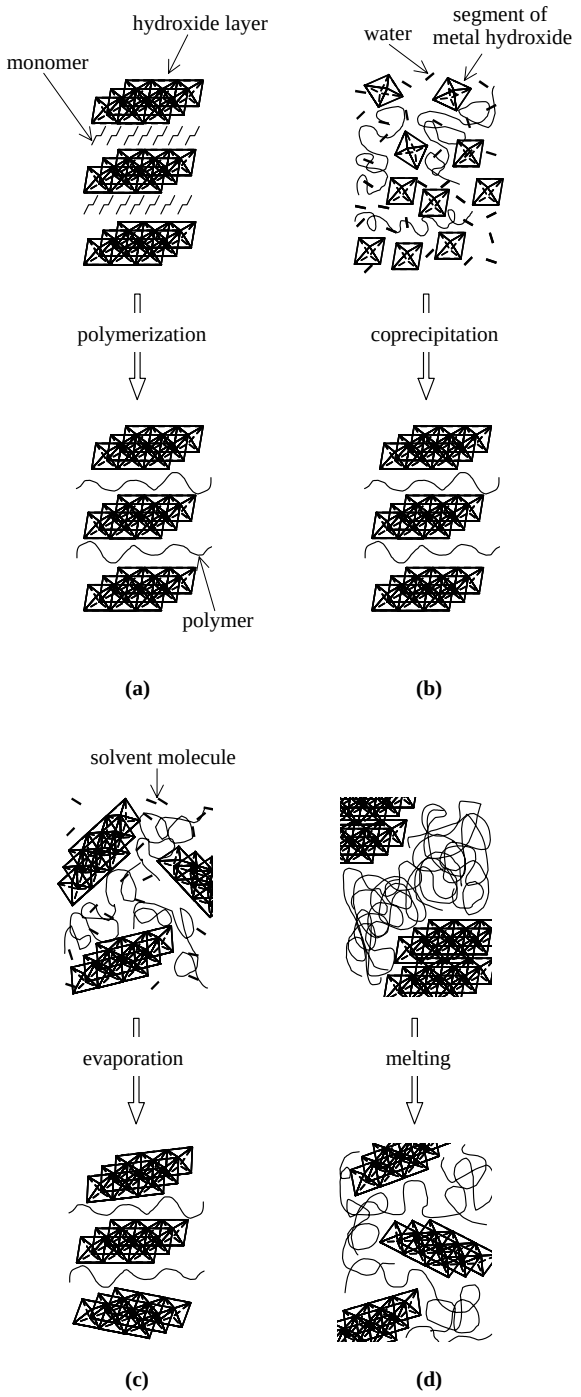


Figure 7. Synthetic routes to polymer/lamellar solid nanocomposites: (a) *in-situ* polymerization method, (b) template method, (c) delamination–reaggregation method, and (d) melt intercalation method.

phase mixing occurs on a macroscopic scale, and the inorganic additives are not homogeneously dispersed at the microscopic level (Fig. 8a). In contrast, nanocomposites are a new class of materials with ultrafine phase dimensions on the nanometer scale, and owing to intimate mixing and improved interfacial properties, they perform better than conventional composites.

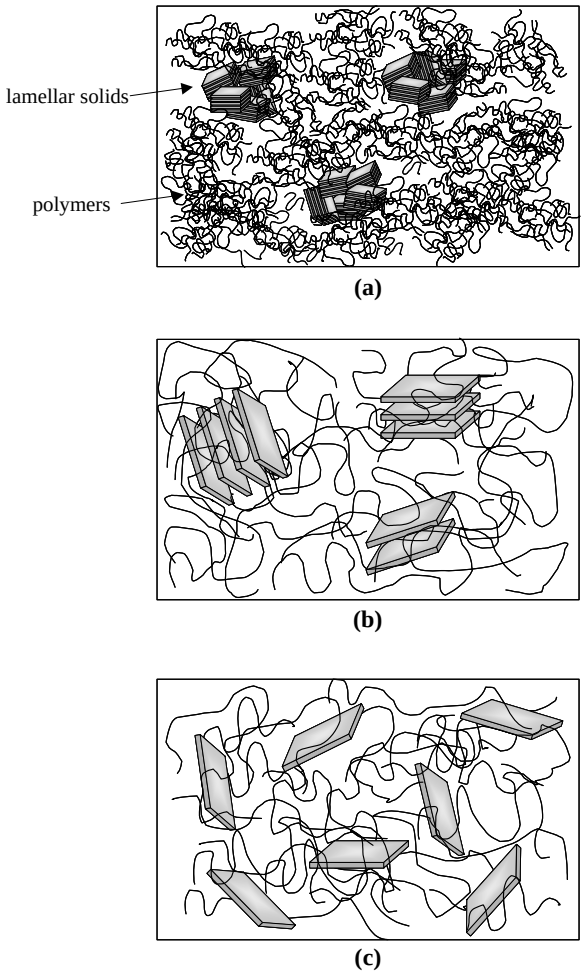


Figure 8. Types of polymer/lamellar solid composites: (a) conventional composite, (b) intercalated nanocomposite, and (c) delaminated nanocomposite. Panel a is drawn on a reduced scale relative to panels b and c.

From a structural viewpoint, there are two types of nanocomposites of lamellar compounds. In intercalated hybrids, a single polymer layer or a few layers are intercalated into the interlayer galleries (Fig. 8b). In delaminated hybrids, single layers of lamellar solids are uniformly dispersed in polymer matrices (Fig. 8c).

Historically, LDHs have been used as flame retardants and as acid neutralizers in stabilizer packages for polymers [97–100]. In these cases, polymers and LDHs form conventional composites. As for nanocomposites of LDHs, they are still at the experimental stage. Most of the polymer/LDH nanocomposite materials reported to date can be categorized as intercalated hybrids. Chemical compositions, basal spacings, and synthetic methods for polymer/LDH nanocomposite materials are summarized in Table 2.

The *in-situ* polymerization method is widely employed for various lamellar materials, such as smectite clays [101–111], graphite oxides [112], zirconium phosphates [113], $RuCl_3$ [114], $HNbMoO_6$ [115], MoO_3 [116, 117], V_2O_5 [118, 119], and FeOCl [120, 121]. Several studies related to LDHs have been also reported; polyacrylate, polyaniline, poly(aminobenzoate), and poly(α,β-aspartate) have been

Table 2. Polymer/LDH nanocomposite materials.

Composition of host layer	Polymer	Basal spacing / Å	Method	Ref.
Mg–Al	polyacrylate	13.4	*in-situ*	[56]
Mg–Al	polyacrylate	poorly crystalline	*in-situ*	[93]
Mg–Al	polystyrene oligomer	23.2	ion exchange	[94]
Li–Al	poly(aminobenzoate)	15.5	*in-situ*	[122]
Ni–Fe	polyacrylate	12.6	template	[125]
Ni–Fe	polyacrylate	12.6	*in-situ*	[125]
Cu–Cr	polyaniline	13.3	*in-situ*	[126]
Cu–Al	polyaniline	13.5	*in-situ*	[126]
Mg–Al	poly(α, β-aspartate)	9.0	*in-situ*	[127]
Mg–Al	poly(α, β-aspartate)	15.1	template	[127]
Mg–Al	poly(styrenesulfonate)	21.0	template	[131]
Mg–Al	poly(vinyl sulfonate)	12.7	template	[131]
Ca–Al	poly(vinyl alcohol)	18.0	template	[132]
Mg–Al	poly(styrenesulfonate)	20.8	template	[133]
Zn–Al	poly(styrenesulfonate)	21.6	template	[133]
Mg–Al	poly(styrenesulfonate)	20.8	template	[134]
Mg–Al	poly(vinyl sulfonate)	13.1	template	[134]
Mg–Al	polyacrylate	12.0	template	[134]
Ca–Al	poly(styrenesulfonate)	19.6	template	[134]
Ca–Al	poly(vinyl sulfonate)	13.2	template	[134]
Ca–Al	polyacrylate	12.4	template	[134]
Zn–Al	poly(styrenesulfonate)	21.6	template	[134]
Zn–Al	poly(vinyl sulfonate)	13.3, 13.0	template	[134]
Zn–Al	polyacrylate	12.4	template	[134]
Co–Al	poly(vinyl sulfonate)	13.3	template	[134]
Zn–Al	poly(styrenesulfonate)	poorly crystalline	delamination–reaggregation	[135]

intercalated into LDHs by *in-situ* polymerization [56, 93, 122–127]. LDHs intercalated with monomers are generally obtained in crystalline phases, and polymerization is usually initiated thermally. Transformation from monomers to polymers is often confirmed by IR analysis [56, 122, 125, 126]. XRD analysis is not suitable to confirm the polymerization because the basal spacing changes are very small.

Other intriguing uses of LDHs have been also reported related to *in-situ* polymerization method [94]. During the polymerization of styrene in aqueous emulsions, LDHs have been added to trap the anionic oligomers for the purpose of monitoring the intermediate states. Moreover, investigation of photoinduced isomerization and polymerization of organic guests is another interesting study [128]. Polymerization and isomerization of muconate anions have been controlled by media used for suspension.

The template method is an advantageous technique for use with LDHs. (There have been only a few reports about the use of this method for lamellar compounds other than LDHs [129, 130].) Poly(α, β-aspartate), poly(styrenesulfonate), polyacrylate, poly(vinyl sulfonate), and poly(vinyl alcohol) have been intercalated into LDH hosts by means of coprecipitation or soft chemistry preparation [125, 127, 131–134]. This method can be used only with water-soluble polymers. The capture of polymer chains by means of the memory effect could also be categorized as a template method [135]. Calcined LDHs have been reconstructed in the presence of anionic polymers, and ordered hybrids have been obtained.

The delamination–reaggregation method is new to LDHs. Delamination followed by restacking has been investigated recently [136]. A possible approach to the fabrication of polymer/LDH nanocomposites via a delamination and reaggregation process has also been reported [135]. The nanocomposite obtained via this approach shows a poorly crystalline phase. Numerous nanocomposites composed of polymers and other lamellar compounds, such as smectite clays, MoO_3, and graphite oxides, have been obtained by this method, and the products are usually intercalated hybrids rather than delaminated hybrids [8–18].

Finally, the melt intercalation method is a practical one because a conventional polymer extrusion or injection molding process can be used. This method has been established for smectite clays with polymers of nylon 6, epoxy, rubbers, and polyolefins [7, 95, 103, 137–147]. Although this method has not yet been applied to LDH nanocomposites, reactions of carbonate LDHs with molten organic acids, such as sebacic, capric, decanoic, and phenylphosphonic acids, have been tested, and it has been found that the organic moieties are partly inserted into the interlayer regions [148–151].

3.3. Nanocomposite Materials from Layered Double Hydroxides

In this section, we review other nanocomposite materials made from LDHs, focusing on the immobilization of guest molecules in the interlayers of LDHs. Through anion

exchange or memory effect processes, a wide range of functional materials can be encapsulated within the nanometer-scale texture of LDHs. The charge density and chemical composition of LDHs can easily be controlled by means of the simple synthetic methods used to prepare them. Control of the charge density permits regulation of the number of guest anions, and compositional diversity can be used to impart electrochemical or redox properties to LDHs. In combination with intercalation reactions, and control of the chemical composition of LDHs should lead to a variety of applications for LDH-based nanocomposites.

LDHs have been used to immobilize biomimetic catalysts [152–155] and photocatalysts [156]. Porphyrins or phthalocyanines, an example of the former, have been confirmed to serve as homogeneous catalysts in aqueous solutions. However, the use of homogeneous catalysts in environmental remediation technologies is impractical because separation of the aqueous catalysts, reactants, and products is difficult. Porphyrins or phthalocyanines intercalated into LDHs by anion exchange, memory effect, and template methods have been verified to be active and separable from the reaction solutions, although some inevitable drop in activity is observed, owing to hindered access to intercalated porphyrins or phthalocyanines.

Polyoxometalate photocatalysts, such as $[W_7O_{24}]^{6-}$, $[W_{10}O_{32}]^{4-}$, $[PW_{12}O_{40}]^{3-}$, $[SiW_{12}O_{40}]^{4-}$, and $[P_2W_{18}O_{62}]^{6-}$, effectively degrade organochlorine pesticides. Again, these polyoxometalates are water soluble, and water-insoluble supports are necessary for recycling of the catalysts. LDHs intercalated with $[W_7O_{24}]^{6-}$ anions have been synthesized via anion exchange, and they show photocatalytic activity against hexachlorocyclohexane.

The creation of novel microporous materials by means of intercalation of polyoxometalates into LDHs has also been intensively studied [19, 47–49, 157, 158]. These polyoxometalate- LDHs are synthesized by means of anion exchange, memory effect, and template methods. However, a delamination method like the one used to intercalate bulky inorganic ions into zirconium phosphates [20, 21, 67, 68] might be a better method for intercalation of polyoxometalates into LDHs.

Immobilization of electroactive anions in LDHs has been also reported. LDHs containing transition metals such as Ni, Co, and Mn have been found to be redox active [159, 160]. When these LDHs are used as electrode modifiers, they immobilize electroactive anions in the interlayers and enhance the electron-transfer rate toward the intercalated ions. Thin hybrid films of an amphiphilic anionic Ru(II) complex with Ni–Al LDHs have been prepared by the Langmuir–Blodgett (LB) method [161]. The nickel sites in the LDH layers act as mediators, oxidizing the Ru(II) cyano complex in the interlayers. The hybrid films may thus be used as electrochemical sensors of, for example, bimolecular materials. However, these hybrid films are composed of 30–70 LDH layers, not single layers, because LDHs are not delaminated under the preparation conditions used. In contrast, layer-by-layer assemblies have been developed for other lamellar compounds. Single layers of titania or zirconium phosphate and polycations have been alternately deposited on substrates [17, 18], and LB films prepared with delaminated smectite clay layers have been also reported [162]. The delamination techniques recently elaborated for LDHs may also provide a method for fabricating well-structured thin LDH hybrid films.

There have also been reports of the incorporation of C_{60} into LDHs by anion exchange with C_{60} charged by chemical reduction or by co-intercalation with organic anions and solvents [163, 164]. The interaction between LDH layers and C_{60} alters the photophysical properties of C_{60} and enhances its photoluminescence. There have also been attempts to intercalate a range of biochemical substances into LDHs, mainly by anion exchange, with the aim of developing pharmaceutical or cosmetic uses [165–172]. Amino acids, vitamins, enzymes, nucleosides, and adenosine triphosphate (ATP) have reportedly been used for this purpose. The modification of LDHs by organic surfactants and various organic anions has been intensively studied with the goal of changing the surface properties of the LDHs [173–176]. In addition, such organic anion intercalates of LDHs have been subjected to calcination, which produces carbon/LDH nanocomposites, and the catalytic activity or porosity of these composites has been evaluated [177–179]. Polymerization of silicates in LDHs has been also studied [180, 181] and appears to be a new synthetic approach to the preparation of smectite-like clays via a chimie duce route.

4. SUMMARY

LDHs are versatile compounds for fabricating nanocomposites with a wide range of materials. The lamellar nature of LDHs permits host–guest chemistry and intercalation reactions, which invoke considerable attention from material designers.

Delamination technique is one of the most versatile routes to fabricate nanocomposites. The novel delamination techniques for LDHs that have been reported should trigger the development of a variety of new nanocomposites of LDHs. Moreover, the synthetic route of LDHs is environmentally benign, which has become increasingly important recently. There are many other valuable features of LDHs, such as facileness in synthesis, anion exchange ability, versatility in chemical composition, and memory effect as an intercalation method. The number of reports about LDHs has considerably increased during the past decade, and the above-mentioned useful characteristics of LDHs for fabrication of new functional materials including nanocomposites should stimulate even more research.

GLOSSARY

Delamination Phenomenon that layers of lamellar solids come apart into single layers or packets of a few layers in appropriate solutions or media.

Exfoliation Has the same meaning as delamination.

Hydrotalcite-like compound Other general term for layered double hydroxide.

Intercalation Insertion of molecules or ions into interlayers in lamellar solid.

Interlayer The weakly interacting region between the stacked sheets in lamellar solid.

Lamellar solids Compounds that possess layered structure.

Layered double hydroxides A class of lamellar solids, consisting of positively charged hydroxide layers with charge-balancing anions and water molecules in interlayers.

Layered structure The structure in which basal sheets of atoms that are tightly bonded within the sheets but relatively weakly bonded in the direction perpendicular to the sheets are stacked on top of one another.

REFERENCES

1. Ph. Refait, M. Abdelmoula, and J.-M. R. Génin, *Corros. Sci.* 40, 1547 (1998).
2. H. C. B. Hansen and I. F. Poulsen, *Clays Clay Miner.* 47, 312 (1999).
3. A. G. B. Williams and M. M. Scherer, *Environ. Sci. Technol.* 35, 3488 (2001).
4. F. Cavani, F. Trifirò, and A. Vaccari, *Catal. Today* 11, 173 (1991).
5. A. Vaccari, *Catal. Today* 41, 53 (1998).
6. A. Vaccari, *Appl. Clay Sci.* 14, 161 (1999).
7. A. Okada and A. Usuki, *Mater. Sci. Eng. C* 3, 109 (1995).
8. H.-L. Tsai, J. L. Schindler, C. R. Kannewurf, and M. G. Kanatzidis, *Chem. Mater.* 9, 875 (1997).
9. Y. Matsuo, K. Hatase, and Y. Sugie, *Chem. Mater.* 10, 2266 (1998).
10. Y. Matsuo, K. Tahara, and Y. Sugie, *Carbon* 35, 113 (1997).
11. M. G. Kanatzidis, R. Bissessur, D. C. DeGroot, J. L. Schindler, and C. R. Kannewurf, *Chem. Mater.* 5, 595 (1993).
12. L. F. Nazar, H. Wu, and W. P. Power, *J. Mater. Chem.* 5, 1985 (1995).
13. L. Wang, J. Schindler, C. R. Kannewurf, and M. G. Kanatzidis, *J. Mater. Chem.* 7, 1277 (1997).
14. G. Lagaly, *Appl. Clay Sci.* 15, 1 (1999).
15. R. Schöllhorn, *Chem. Mater.* 8, 1747 (1996).
16. N. Sukpirom and M. M. Lerner, *Chem. Mater.* 13, 2179 (2001).
17. T. Sasaki, Y. Ebina, T. Tanaka, M. Harada, M. Watanabe, and G. Decher, *Chem. Mater.* 13, 4661 (2001).
18. H.-N. Kim, S. W. Keller, T. E. Mallouk, J. Schmitt, and G. Decher, *Chem. Mater.* 9, 1414 (1997).
19. T. J. Pinnavia, *Adv. Chem. Ser.* 245, 283 (1995).
20. E. Rodríguez-Castellón, A. Jiménez-López, P. Maireles-Torres, J. Mérida-Robles, P. Braos-García, G. Aguilar-Armenta, E. Flores-Loyola, F. Marmonttini, and E. Felici, *Langmuir* 17, 3769 (2001).
21. G. Alberti, S. Cavalaglio, F. Marmottini, K. Matusek, J. Megyeri, and L. Szirtes, *Appl. Catal. A-Gen.* 218, 219 (2001).
22. T. Sumida, R. Abe, M. Hara, J. N. Kondo, and K. Domen, *J. Mater. Res.* 15, 2587 (2000).
23. R. Toranzo, M. A. Vicente, M. A. Bañares-Muñoz, L. M. Gandía, and A. Gil, *Micropor. Mesopor. Mater.* 173, 173 (1998).
24. M. Adachi-Pagano, C. Forano, and J.-P. Besse, *Chem. Commun.* 91 (2000).
25. L. Ingram and H. F. W. Taylor, *Mineral. Mag.* 36, 465 (1967).
26. R. Allmann, *Acta Crystallogr.* B24, 972 (1968).
27. C. J. Serna, J. L. Rendon, and J. E. Iglesias, *Clays Clay Miner.* 30, 180 (1982).
28. A. M. Fogg, J. S. Dunn, and D. O'Hare, *Chem. Mater.* 10, 356 (1998).
29. N. Iyi, K. Kurashima, and T. Fujita, *Chem. Mater.* 14, 583 (2002).
30. S. P. Newman and W. Jones, *New J. Chem.* 22, 105 (1998).
31. S. P. Newman and W. Jones, in "Supramolecular Organization and Materials Design," (W. Jones and C. N. R. Rao, Eds.), pp. 295–331. Cambridge University Press, Cambridge, 2001.
32. S. Miyata, *Clays Clay Miner.* 28, 50 (1980).
33. G. J. Ross and H. Komada, *Am. Mineral.* 52, 1036 (1967).
34. J. W. Boclair and P. S. Braterman, *Chem. Mater.* 10, 2050 (1998).
35. S. Miyata, *Clays Clay Miner.* 31, 305 (1983).
36. J. Inacio, C. Taviot-Guého, C. Forano, and J. P. Besse, *Appl. Clay Sci.* 18, 255 (2001).
37. M. C. Hermosín, I. Pavlovic, M. A. Ulibarri, and J. Cornejo, *Water Resour.* 30, 171 (1996).
38. D. Tichit, M. N. Bennani, F. Figueras, and J. R. Ruiz, *Langmuir* 14, 2086 (1998).
39. T. Sato, K. Kato, T. Endo, and M. Shimada, *React Solids* 2, 253 (1986).
40. L. Pesic, S. Salipurovic, V. Markovic, D. Vucelic, W. Kagunya, and W. Jones, *J. Mater. Chem.* 2, 1069 (1992).
41. F. Rey, V. Fornés, and J. M. Rojo, *J. Chem. Soc., Faraday Trans.* 88, 2233 (1992).
42. K. J. D. MacKenzie, R. H. Meinhold, B. L. Scherriff, and Z. Xu, *J. Mater. Chem.* 3, 1263 (1993).
43. T. Hibino, Y. Yamashita, K. Kosuge, and A. Tsunashima, *Clays Clay Miner.* 43, 427 (1995).
44. M. J. Hudson, S. Carlino, and D. C. Apperley, *J. Mater. Chem.* 5, 323 (1995).
45. W. T. Reichle, S. Y. Kang, and D. S. Everhardt, *J. Catal.* 101, 352 (1986).
46. K. Chibwe and W. Jones, *J. Chem. Soc., Chem. Commun.* 926 (1989).
47. K. Chibwe and W. Jones, *Chem. Mater.* 1, 489 (1989).
48. E. D. Dimotakis and T. J. Pinnavia, *Inorg. Chem.* 29, 2393 (1990).
49. E. Narita, P. Kaviratna, and T. J. Pinnavaia, *Chem. Lett.* 805 (1991).
50. F. Kooli, C. Depège, A. Ennaqadi, A. de Roy, and J. P. Besse, *Clays Clay Miner.* 45, 92 (1997).
51. T. Hibino and A. Tsunashima, *J. Mater. Sci. Lett.* 19, 1403 (2000).
52. T. Sato, H. Fujita, T. Endo, M. Shimada, and A. Tsunashima, *React. Solids* 5, 219 (1988).
53. B. Rebours, J.-B. E. de la Caillerie, and O. Clause, *J. Am. Chem. Soc.* 116, 1707 (1994).
54. L. M. Parker, N. B. Milestone, and R. H. Newman, *Ind. Eng. Chem. Res.* 34, 1196 (1995).
55. M. A. Ulibarri, I. Pavlovic, C. Barriga, M. C. Hermosín, and J. Cornejo, *Appl. Clay Sci.* 18, 17 (2001).
56. M. Tanaka, I. Y. Park, K. Kuroda, and C. Kato, *B. Chem. Soc. Jpn.* 62, 3442 (1989).
57. H.-S. Shin, M.-J. Kim, S.-Y. Nam, and H.-C. Moon, *Water Sci. Technol.* 34, 161 (1996).
58. I. C. Chisem, S. D. Cosgrove, and W. Jones, *J. Therm. Anal.* 50, 757 (1997).
59. Y. Fukushima, *Clays Clay Miner.* 32, 320 (1984).
60. K. Norrish, in "Coagulation and Flocculation, Discussions of the Faraday Society," (Faraday Society, Ed.), no. 18, pp. 120–134. Aberdeen University Press, London, 1954.
61. H. van Olphen, "An Introduction to Clay Colloid Chemistry: For Clay Technologists, Geologists, and Soil Scientists," pp. 150–161. Wiley, New York, 1977.
62. P. Liu and K. Gong, *Carbon* 37, 706 (1999).
63. I. Bezverkhy, P. Afanasiev, and M. Lacroix, *Inorg. Chem.* 39, 5416 (2000).
64. J. Heising and M. G. Kanatzidis, *J. Am. Chem. Soc.* 121, 638 (1999).
65. P. Joensen, E. D. Crozier, N. Alberding, and R. F. Frindt, *J. Phys. C Solid State Phys.* 20, 4043 (1987).
66. D. M. Kaschak, S. A. Johnson, D. E. Hooks, H.-N. Kim, M. D. Ward, and T. E. Mallouk, *J. Am. Chem. Soc.* 120, 10887 (1998).
67. J. Xu and Z. Gao, *Micropor. Mesopor. Mater.* 24, 213 (1998).
68. F. J. Pérez-Reina, P. Olivera-Pastor, P. Maireles-Torres, E. Rodríguez-Castellón, and A. Jiménez-López, *Langmuir* 14, 4017 (1998).
69. G. Alberti, E. Giontella, and S. Murcia-Mascarós, *Inorg. Chem.* 36, 2844 (1997).
70. G. Alberti, C. Dionigi, E. Giontella, S. Murcia-Mascarós, and R. Vivani, *J. Colloid Interface Sci.* 188, 27 (1997).

71. J. Xu, Y. Tang, H. Zhang, and Z. Gao, *J. Inclusion Phenom. Mol. Recognit. Chem.* 27, 303 (1997).
72. K. Peeters, P. Grobet, and E. F. Vansant, *J. Mater. Chem.* 6, 239 (1996).
73. P. Maireles-Torres, P. Olivera-Pastor, E. Rodriguez-Castellon, A. Jimenez-Lopez, and A. A. G. Tomlinson, *J. Mater. Chem.* 1, 739 (1991).
74. G. Alberti, M. Casciola, and U. Costantino, *J. Colloid Interface Sci.* 107, 256 (1985).
75. G. Alberti, S. Cavalaglio, C. Dionigi, and F. Marmottini, *Langmuir* 16, 7663 (2000).
76. T. Nakato, Y. Furumi, N. Terao, and T. Okuhara, *J. Mater. Chem.* 10, 737 (2000).
77. Y. Matsumoto, A. Funatsu, D. Matsuo, U. Unal, and K. Ozawa, *J. Phys. Chem. B* 105, 10893 (2001).
78. T. Sasaki, Y. Ebina, Y. Kitami, M. Watanabe, and T. Oikawa, *J. Phys. Chem. B* 105, 6116 (2001).
79. R. Abe, S. Ikeda, J. N. Kondo, M. Hara, and K. Domen, *Thin Solid Films* 344, 156 (1999).
80. T. Sasaki, *Supramol. Sci.* 5, 367 (1998).
81. R. Abe, K. Shinohara, A. Tanaka, M. Hara, J. N. Kondo, and K. Domen, *Chem. Mater.* 10, 329 (1998).
82. T. Sasaki and M. Watanabe, *J. Am. Chem. Soc.* 120, 4682 (1998).
83. T. Sasaki, S. Nakano, S. Yamauchi, and M. Watanabe, *Chem. Mater.* 9, 602 (1997).
84. T. Sasaki, M. Watanabe, H. Hashizume, H. Yamada, and H. Nakazawa, *J. Am. Chem. Soc.* 118, 8329 (1996).
85. R. E. Schaak and T. E. Mallouk, *Chem. Commun.* 706 (2002).
86. J.-Y. Kim, I. Chung, and J.-H. Choy, *Chem. Mater.* 13, 2759 (2001).
87. Y.-S. Han, I. Park, and J.-H. Choy, *J. Mater. Chem.* 11, 1277 (2001).
88. R. E. Schaak and T. E. Mallouk, *Chem. Mater.* 12, 3427 (2000).
89. G. B. Saupe, C. C. Waraksa, H.-N. Kim, Y. J. Han, D. M. Kaschak, D. M. Skinner, and T. E. Mallouk, *Chem. Mater.* 12, 1556 (2000).
90. L. F. Braganza, R. J. Crawford, M. V. Smalley, and R. K. Thomas, *Clays Clay Miner.* 38, 90 (1990).
91. A. I. Bortun, S. A. Khainakov, L. N. Bortun, E. Jaimez, J. R. García, and A. Clearfield, *Mater. Res. Bull.* 34, 921 (1999).
92. T. Hibino and W. Jones, *J. Mater. Chem.* 11, 1321 (2001).
93. S. O'Leary, D. O'Hare, and G. Seeley, *Chem. Commun.* 1506 (2002).
94. G. T. D. Shouldice, P. Y. Choi, B. E. Koene, L. F. Nazar, and A. Rudin, *J. Polym. Sci., Polym. Chem.* 33, 1409 (1995).
95. E. P. Giannelis, *Adv. Mater.* 8, 29 (1996).
96. M. Alexandre and P. Dubois, *Mater. Sci. Eng.* 28, 1 (2000).
97. G. Camino, A. Maffezzoli, M. Braglia, M. De Lazzaro, and M. Zammarano, *Polym. Degrad. Stabil.* 74, 457 (2001).
98. S. H. Patel, M. Xanthos, J. Grenci, and P. B. Klepak, *J. Vinyl Addit. Technol.* 1, 201 (1995).
99. L. van der Ven, M. L. M. van Gemert, L. F. Batenburg, J. J. Keern, L. H. Gielgens, T. P. M. Koster, and H. R. Fischer, *Appl. Clay Sci.* 17, 25 (2000).
100. K. Mori, Y. Nakamura, and I. Kikuchi, *J. Polym. Sci., Polym. Lett. Ed.* 19, 623 (1981).
101. M. P. Eastman, E. Bain, T. L. Porter, K. Manygoats, R. Whitehorse, R. A. Parnell, and M. E. Hagerman, *Appl. Clay Sci.* 15, 173 (1999).
102. M. Alexandre, P. Dubois, T. Sun, J. M. Garces, and R. Jérôme, *Polymer* 43, 2123 (2002).
103. P. C. LeBaron, Z. Wang, and T. J. Pinnavaia, *Appl. Clay Sci.* 15, 11 (1999).
104. Z. Wang and T. J. Pinnavaia, *Chem. Mater.* 10, 1820 (1998).
105. H. Shi, T. Lan, and T. J. Pinnavaia, *Chem. Mater.* 8, 1584 (1996).
106. A. Usuki, M. Kawasumi, Y. Kojima, A. Okada, T. Kurauchi, and O. Kamigaito, *J. Mater. Res.* 8, 1174 (1993).
107. Y. Kojima, A. Usuki, M. Kawasumi, A. Okada, Y. Fukushima, T. Kurauchi, and O. Kamigaito, *J. Mater. Res.* 8, 1185 (1993).
108. A. Usuki, Y. Kojima, M. Kawasumi, A. Okada, Y. Fukushima, T. Kurauchi, and O. Kamigaito, *J. Mater. Res.* 8, 1179 (1993).
109. Y. Fukushima, A. Okada, M. Kawasumi, T. Kurauchi, and O. Kamigaito, *Clay Miner.* 23, 27 (1988).
110. Y. Fukushima and S. Inagaki, *J. Inclusion Phenom.* 5, 473 (1987).
111. T. Lan and T. J. Pinnavaia, *Chem. Mater.* 6, 2216 (1994).
112. P. Liu, K. Gong, P. Xiao, and M. Xiao, *J. Mater. Chem.* 10, 933 (2000).
113. Y. Ding, D. J. Jones, P. Maireles-Torres, and J. Rozière, *Chem. Mater.* 7, 562 (1995).
114. L. Wang, P. Brazis, M. Rocci, C. R. Kannewurf, and M. G. Kanatzidis, *Chem. Mater.* 10, 3298 (1998).
115. H.-J. Nam, H. Kim, S. H. Chang, S.-G. Kang, and S.-H. Byeon, *Solid State Ionics* 120, 189 (1999).
116. R. Bissessur, D. C. DeGroot, J. L. Schindler, C. R. Kannewurf, and M. G. Kanatzidis, *J. Chem. Soc., Chem. Commun.* 687 (1993).
117. T. A. Kerr, H. Wu, and L. F. Nazar, *Chem. Mater.* 8, 2005 (1996).
118. M. G. Kanatzidis, C.-G. Wu, H. O. Marcy, D. C. DeGroot, and C. R. Kannewurf, *Chem. Mater.* 2, 222 (1990).
119. Y.-J. Liu, D. C. DeGroot, J. L. Schindler, C. R. Kannewurf, and M. G. Kanatzidis, *J. Chem. Soc., Chem. Commun.* 593 (1993).
120. C.-G. Wu, D. C. DeGroot, H. O. Marcy, J. L. Schindler, C. R. Kannewurf, T. Bakas, V. Papaefthymiou, W. Hirpo, J. P. Yesinowski, Y.-J. Liu, and M. G. Kanatzidis, *J. Am. Chem. Soc.* 117, 9229 (1995).
121. M. G. Kanatzidis, L. M. Tonge, T. J. Marks, H. O. Marcy, and C. R. Kannewurf, *J. Am. Chem. Soc.* 109, 3797 (1987).
122. V. P. Isupov, L. E. Chupakhina, M. A. Ozerova, V. G. Kostrovsky, and V. A. Poluboyarov, *Solid State Ionics* 141–142, 231 (2001).
123. T. Shichi, S. Yamashita, and K. Takagi, *Supramol. Sci.* 5, 303 (1998).
124. S. Katahira, K. Yasue, and M. Inagaki, *J. Mater. Res.* 14, 1178 (1999).
125. S. Rey, J. Mérida-Robles, K.-S. Han, L. Guerlou-Demourgues, C. Delmas, and E. Duguet, *Polym. Int.* 48, 277 (1999).
126. T. Challier and R. C. T. Slade, *J. Mater. Chem.* 4, 367 (1994).
127. N. T. Whilton, P. J. Vickers, and S. Mann, *J. Mater. Chem.* 7, 1623 (1997).
128. S. W. Rhee and D.-Y. Jung, *Bull. Korean Chem. Soc.* 23, 35 (2002).
129. K. A. Carrado, P. Thiyagarajan, and D. L. Elder, *Clays Clay Miner.* 44, 506 (1996).
130. K. A. Carrado and L. Xu, *Chem. Mater.* 10, 1440 (1998).
131. O. C. Wilson Jr., T. Olorunyolemi, A. Jaworski, L. Borum, D. Young, A. Siriwat, E. Dickens, C. Oriakhi, and M. Lerner, *Appl. Clay Sci.* 15, 265 (1999).
132. P. B. Messersmith and S. I. Stupp, *Chem. Mater.* 7, 454 (1995).
133. C. O. Oriakhi, I. V. Faar, and M. M. Lerner, *Clays Clay Miner.* 45, 194 (1997).
134. C. O. Oriakhi, I. V. Faar, and M. M. Lerner, *J. Mater. Chem.* 6, 103 (1996).
135. F. Leroux and J.-P. Besse, *Chem. Mater.* 13, 3507 (2001).
136. F. Leroux, M. Adachi-Pagano, M. Intissar, S. Chauvière, C. Forano, and J.-P. Besse, *J. Mater. Chem.* 11, 105 (2001).
137. J. W. Cho and D. R. Paul, *Polymer* 42, 1083 (2001).
138. P. H. Nam, P. Maiti, M. Okamoto, T. Kotaka, N. Hasegawa, and A. Usuki, *Polymer* 42, 9633 (2001).
139. N. Hasegawa, H. Okamoto, M. Kawasumi, and A. Usuki, *J. Appl. Polym. Sci.* 74, 3359 (1999).
140. L. Liu, Z. Qi, and X. Zhu, *J. Appl. Polym. Sci.* 71, 1133 (1999).
141. M. Kato, A. Usuki, and A. Okada, *J. Appl. Polym. Sci.* 66, 1781 (1997).
142. R. A. Vaia and E. P. Giannelis, *Macromolecules* 30, 8000 (1997).
143. R. A. Vaia and E. P. Giannelis, *Macromolecules* 30, 7990 (1997).
144. R. A. Vaia, B. B. Sauer, O. K. Tse, and E. P. Giannelis, *J. Polym. Sci., Polym. Phys.* 35, 59 (1997).

145. R. A. Vaia, K. D. Jandt, E. J. Kramer, and E. P. Giannelis, *Chem. Mater.* 8, 2628 (1996).
146. R. A. Vaia, S. Vasudevan, W. Krawiec, L. G. Scanlon, and E. P. Giannelis, *Adv. Mater.* 7, 154 (1995).
147. R. A. Vaia, H. Ishii, and E. P. Giannelis, *Chem. Mater.* 5, 1694 (1993).
148. S. Carlino, *Solid State Ionics* 98, 73 (1997).
149. S. Carlino, M. J. Hudson, S. W. Husain, and J. A. Knowles, *Solid State Ionics* 84, 117 (1996).
150. S. Carlino and M. J. Hudson, *J. Mater. Chem.* 5, 1433 (1995).
151. S. Carlino and M. J. Hudson, *J. Mater. Chem.* 4, 99 (1994).
152. L. Ukrainczyk, M. Chibwe, T. J. Pinnavaia, and S. A. Boyd, *Environ. Sci. Technol.* 29, 439 (1995).
153. S. Bonnet, C. Forano, A. de Roy, and J. P. Besse, *Chem. Mater.* 8, 1962 (1996).
154. B. Sels, D. De Vos, M. Buntinx, F. Pierard, A. Kirsch-De Mesmaeker, and P. Jacobs, *Nature* 400, 855 (1999).
155. M. Chibwe, L. Ukrainczyk, S. A. Boyd, and T. J. Pinnavaia, *J. Mol. Catal. A-Chem.* 113, 249 (1996).
156. Y. Guo, D. Li, C. Hu, Y. Wang, E. Wang, Y. Zhou, and S. Feng, *Appl. Catal. B-Environ.* 30, 337 (2001).
157. M. A. Drezdzon, *Inorg. Chem.* 27, 4628 (1988).
158. T. Hibino and A. Tsunashima, *Chem. Mater.* 9, 2082 (1997).
159. J. Qiu and G. Villemure, *J. Electroanal. Chem.* 395, 159 (1995).
160. J. Qiu and G. Villemure, *J. Electroanal. Chem.* 428, 165 (1997).
161. J. X. He, K. Kobayashi, M. Takahashi, G. Villemure, and A. Yamagishi, *Thin Solid Films* 397, 255 (2001).
162. K. Tamura, H. Setsuda, M. Taniguchi, and A. Yamagishi, *Langmuir* 15, 6915 (1999).
163. W. Ding, G. Gu, W. Zhong, W.-C. Zang, and Y. Du, *Chem. Phys. Lett.* 262, 259 (1996).
164. W.-Y. Tseng, J.-T. Lin, C.-Y. Mou, S. Cheng, S.-B. Liu, P. P. Chu, and H.-W. Liu, *J. Am. Chem. Soc.* 118, 4411 (1996).
165. S.-H. Hwang, Y.-S. Han, and J.-H. Choy, *Bull. Korean Chem. Soc.* 22, 1019 (2001).
166. A. I. Khan, L. Lei, A. J. Norquist, and D. O'Hare, *Chem. Commun.* 2342 (2001).
167. L. Ren, J. He, D. G. Evans, X. Duan, and R. Ma, *J. Mol. Catal. B-Enzym.* 16, 65 (2001).
168. B. Lotsch, F. Millange, R. I. Walton, and D. O'Hare, *Solid State Sci.* 3, 883 (2001).
169. J.-H. Choy, S.-Y. Kwak, J.-S. Park, and Y.-J. Jeong, *J. Mater. Chem.* 11, 1671 (2001).
170. S. Aisawa, S. Takahashi, W. Ogasawara, Y. Umetsu, and E. Narita, *J. Solid State Chem.* 162, 52 (2001).
171. Á Fudala, I. Pálinkó, B. Hrivnák, and I. Kiricsi, *J. Therm. Anal. Calorim.* 56, 317 (1999).
172. Á Fudala, I. Pálinkó, and I. Kiricsi, *Inorg. Chem.* 38, 4653 (1999).
173. S. W. Rhee, J. H. Lee, and D.-Y. Jung, *J. Colloid Interface Sci.* 245, 349 (2002).
174. M. Z. Hussein and H. K. Seng, *J. Mater. Sci. Lett.* 20, 1787 (2001).
175. T. Kanoh, T. Shichi, and K. Takagi, *Chem. Lett.* 117 (1999).
176. M. Jakupca and P. K. Dutta, *Chem. Mater.* 7, 989 (1995).
177. Z. P. Xu, R. Xu, and H. C. Zeng, *Nano Lett.* 1, 703 (2001).
178. K. Putyera, T. J. Bandosz, J. Jagieo, and J. A. Schwarz, *Carbon* 34, 1559 (1996).
179. K. Putyera, T. J. Bandosz, J. Jagiello, and J. A. Schwarz, *Clays Clay Miner.* 42, 1 (1994).
180. S. K. Yun, V. R. L. Constantino, and T. J. Pinnavaia, *Clays Clay Miner.* 43, 503 (1995).
181. C. Depège, F.-Z. El Metoui, C. Forano, A. de Roy, J. Dupuis, and J.-P. Besse, *Chem. Mater.* 8, 952 (1996).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanostructured Metals and Alloys

Dmitri Valentinovich Louzguine, Akihisa Inoue

Tohoku University, Aoba-Ku, Sendai, Japan

CONTENTS

1. INTRODUCTION

Nanostructured metals or more widely speaking metallic materials with a nanoscale size microstructure are relatively new and prospective types of substances that demonstrate superior properties in a wide range. Among them are materials with exclusively high mechanical, enhanced magnetic, electronic, and optical properties. Such properties were obtained due to a very small structure unit (grain, particle, layer) size ranging from 1 to 100 nm, much smaller than that obtained in conventional metallic materials or alloys. Such a unit can have a crystalline, quasicrystalline, or amorphous structure. Due to a small structure unit size these materials are often in highly nonequilibrium state, although lots of them exhibit remarkable stability against structural coarsening. Current and potential applications of nanocrystalline materials are structural materials, different electronic, magnetic, and optical devices, filters, batteries, displays, biomaterials, and protective coatings.

Nanostructured metallic materials (here also called nanomaterials) can be obtained as a group of separated particles or as a polycrystalline sample in a bulk form. The development of age-hardened Al–Cu alloys, discovered by Wilm in the beginning of the 20th century [1], with nanoscale Guinier–Preston zones, θ'' and θ' phases, can be considered one of the first examples of use of the nanocrystals in materials for practical purpose [2]. Several decades later, synthesis of new materials by tailoring of structure at the submicronic, nanoscale, and atomic or molecular scale has become an important interdisciplinary field.

In this chapter we will emphasize metallic nanostructured materials and will not pay much attention to semiconductors, nonmetallic materials, ceramics, and oxides which, however, also produce a large group of technologically important nanomaterials.

2. PRODUCTION

At an early stage of nanomaterials studies ultrafine particles, for example, Pd with face centered cubic (fcc) structure equal to that of the bulk specimen and a mean size of 2 nm, were prepared by a vacuum evaporation method [3]. Fine gold particles were prepared by a gas evaporation technique [4]. Considerable progress of the processing procedure for obtaining nanomaterials has been achieved in the last decade. At present nanomaterials can be produced using various methods from vapor phase (inert gas condensation [5], vapor deposition, sputtering [6, 7]), liquid phase (rapid solidification [8], electrodeposition [9]), or solid phase (mechanical attrition [10], severe plastic deformation [11, 12], devitrification of glassy alloys [13], pulse plating by nanoscale deposits [14], laser beam treatment of the surface, etc.).

The processing procedure can be separated into single-step and double/multistep processes. The former it preferable if the properties are not worsened. Very important from a practical viewpoint, bulk nanomaterials can be produced in a single-step process by casting, for example, or a double-step process like compaction of the preprepared particles.

The same material can be produced by different techniques. For instance, pure nanocrystalline hexagonal close packed (hcp) Co with a grain size below 15 nm has been produced by inert gas condensation [15] and a ball milling technique [16] respectively. However, the method of production often has a significant influence of the properties of nanomaterials.

Mechanical attrition and severe plastic deformation offer the possibility of production of nanomaterials in commercial quantities. The influence of high energy [17] and low energy [18] ball milling on the crystallite size and lattice strain was studied. However, during mechanical attrition the

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (669–697)

problems of powder contamination from milling device and atmosphere, as well as a problem of powder consolidation without coarsening, are still to be solved. Fundamental studies of the microstructural development in ball-milled pure elements have defined the sequence of events leading to nanocrystalline grains with average grain size from 2 to 20 nm [19, 20]. The grain size is found to decrease with milling time down to a constant value. There is a balance between defect creation and recovery during deformation. At the same time, amorphous alloys can also be prepared by ball milling [21].

Nanostructured metals, and even metalloids, can also be produced by severe plastic deformation. For example, nanostructured Si with a mean grain size of less than 20 nm was obtained by torsion straining [22]. The structure is characterized by a small grain size and a specific structure of grain boundaries. A high level of elastic strain and significant microdistortions of the crystal lattice were also observed.

Deformation mechanisms during mechanical milling and a severe plastic deformation process leading to the formation of ultrafine grains require investigation. It is considered that shear processes requiring crystal rotation are required for nanocrystal formation and disclinations may contribute to this process [23]. Very recently high resolution transmission electron microscopy (HRTEM) observation showed partial disclination dipoles in body centered cubic (bcc) Fe that had undergone severe plastic deformation by mechanical milling [24]. It was shown that the formation and migration of such partial disclinations during deformation allow crystalline grains to rotate and rearrange at the nanometer level. The formation of partial disclination dipoles and other disclination defects facilitating deformation under high stresses can also cause considerable strengthening, owing to the interaction of their elastic stress fields with each other and with grain boundaries.

Nanoparticles of crystalline or quasicrystalline phase can be produced directly from the liquid on rapid solidification by the cooling rate-controlled process [25]. For example, the size of the icosahedral particles homogeneously distributed in Zr–Pt [26] alloy is below 10 nm. The formation of the nanoscale icosahedral phase indicates that the icosahedral short-range order exists in the melt of Zr–Pt binary alloy.

Nanoscale particles of a crystalline or quasicrystalline phase can be easily produced by devitrification of the glassy alloys. This indirect method of production of nanostructure requires formation of the glassy phase in the initial stage and its subsequent full or partial devitrification on heating. However, as has been mentioned such a structure can also be obtained directly during rapid solidification or casting by the control of cooling rate.

If the process involves nucleation and growth then a high nucleation rate and low growth rate of the precipitating phase are required in order to obtain a nanostructure. Such conditions are usually obtained under primary devitrification with long-range diffusion controlled growth [27]. For example, Al-rich amorphous alloys with high Al concentrations above 85 at%, in general, suffer double or multistage devitrification in which the first-stage exothermic reaction is due to the precipitation of fcc-Al and the other exothermic peaks result from the decomposition of the residual amorphous phase to intermetallic compounds. Nevertheless, nanocrystalline structure was also obtained after eutectic [28] and polymorphous [29, 30] devitrification of glassy alloys. Another type of phase transformation in amorphous solid leading to formation of a nanostructure is spinodal decomposition [31, 32].

Al–RE–TM (RE—rare earth metals, TM—transition metal) system glassy alloys, produced in late 1980s in Japan [33, 34] and the United States [35], belong to one of the most important groups of Al-based metallic glasses discovered to date. These alloys were obtained by adding TM to binary Al–RE alloys produced shortly before [36, 37]. Simultaneous alloying with RE and TM caused drastic widening of the composition range of the amorphous single phase. These alloys are usually produced in a ribbon shape by the melt spinning technique [38] or in a powder one by gas atomization [39]. It is important that amorphous alloys in the ternary Al–RE–TM system, for example, Al–Y–Ni, in addition to their high strength possess good bend ductility, that is, show the ability to be bent 180° without fracture [40, 41]. An addition of Co partially replacing Y in the $Al_{85}Y_{10}Ni_5$ [here and elsewhere alloy compositions are given in nominal atomic percents (at%)] increased tensile fracture strength, for example from 920 to 1250 MPa in the case of $Al_{85}Y_8Ni_5Co_2$ metallic glass, without worsening of bend ductility [42]. At the same time, the addition of 2 at% Co to $Al_{85}Y_{10}Ni_5$ alloy also caused drastic increment of the glass-forming ability. Critical ribbon thickness being still less than 1 mm increased from 120 to 900 μm [42]. Moreover, high-strength quaternary $Al_{85}Y_8Ni_5Co_2$ alloy shows one of the widest supercooled liquid regions on heating among Al-based metallic glasses.

However, ease of devitrification of the Al-based glassy samples connected with high density of so-called quenched-in nuclei in some glasses as well as low reduced glass-transition and devitrification temperature in other glasses [43] imposes a limit on the sample's critical thickness below 1 mm. At the same time $Al_{85}Ni_{10}Ce_5$ bulk amorphous samples of high relative density were obtained by warm extrusion of atomized amorphous powder [44].

Different Al–RE–TM glasses containing about 85 at% Al show primary precipitation of Al solid solution (α-Al) nanoparticles on heating [45, 46] with extremely high nucleation rate [47, 48] of above 10^{20} $m^{-3}\,s^{-1}$. Fcc α-Al lattice parameter measurements and atom probe ion field microscopy investigation [49] showed a very low concentration of the alloying elements in nanocrystalline Al in accordance with phase diagrams [50] of Al–RE and Al–TM. Segregation of the RE metal having low trace diffusivity in Al on the α-Al/amorphous phase interface is considered to be one of the most important reasons for the low growth rate of α-Al. Nearly the same effects were observed during formation of the α-Fe nanocrystals [51]. It is also important that primary devitrification requires long-range diffusion that impedes crystal growth [52].

One should mention that the devitrification behavior of various Al–Y–Ni–Co alloys has been also studied, and at definite compositions formation of the α-Al was observed at the primary stage [53, 54]. Y-, Sm-, Gd-, and Dy-bearing metallic glasses of the $Al_{85}RE_8Ni_5Co_2$ composition showed precipitation of the α-Al nanoparticles after continuous

heating using differential scanning calorimetry (DSC) at a high enough heating rate (0.67 K/s and higher) or annealing at the temperature above T_g. For the aforementioned alloys a heating rate of 0.67 K/s is high enough to prevent devitrification below T_g. At the same time, Y-, Gd-, and Dy-bearing metallic glasses [55] as well as $Al_{85}Y_4Nd_4Ni_5Co_2$ [56] showed simultaneous formation of the intermetallic compound(s) and α-Al nanoparticles, or primary formation of the intermetallic compound only after annealing below T_g. For example, the $Al_{85}Y_8Ni_5Co_2$ alloy shows formation of an unknown intermetallic compound conjointly with α-Al nanoparticles after annealing up to the completion of the primary phase transformation. After the completion of the primary phase transformation a lot of relatively small intermetallic compound particles below 50 nm size exist in the structure of the $Al_{85}Y_8Ni_5Co_2$ alloy. The intermetallic compounds are metastable and have a multicomponent composition. The volume fraction of the intermetallic compound is higher than that of α-Al and the fraction of α-Al depends upon annealing temperature below T_g. It should be also noted that the multicomponent intermetallic phase is a leading phase during isothermal devitrification of $Al_{85}Y_4Nd_4Ni_5Co_2$ below T_g (i.e., starts to precipitate prior to α-Al).

It has been also reported that in some Al–TM–RE glasses devitrification appears to be preceded by amorphous phase separation that leads to formation of a nanostructure [57].

Devitrification of nanocrystal-forming Ti-based alloys, for example, the $Ti_{50}Ni_{20}Cu_{23}Sn_7$ alloy, begins from the primary precipitation of a nanoscale equiaxed, almost spherical particles of Ti_2Ni solid solution (other alloying elements are partially dissolved in this phase) with a lattice parameter of 1.138 nm (Fig. 1) [58]. Although only Ti_2Ni solid solution phase was observed to precipitate during the primary devitrification reaction the trace of heat release corresponding to primary devitrification reaction studied by DSC and isothermal calorimetry (Fig. 2) has a complicated shape. This demonstrates a change of the growth rate during primary crystallization according to Kolmogorov [59], Johnson and Mehl [60], and Avrami [61], and Kelton [62]. The general exponential equation for the fraction transformed is

$$x_e(t) = \frac{4\pi}{3V_0}\int_0^t I(\tau)\left[\int_\tau^t g(t')\,dt'\right]^3 d\tau \qquad (1)$$

where $I(\tau)$ and $g(t')$ are time-dependent nucleation and growth rates, respectively. This feature has also been observed in Zr-based alloy [63]. Difficulty of nucleation of the Ti_2Ni solid solution having a large lattice parameter and complicated composition is one of the reasons for the elevated glass-forming ability of the Ti–Ni–Cu–Sn amorphous alloy studied.

The largest Ti_2Ni particle size in the beginning of the transformation and after its completion is plotted versus transformation time in Figure 3. Crystal size (average of three measurements for the largest particle) versus transformation time (without incubation period) dependence has a nonlinear character and indicates a diffusion-controlled growth. Cu atoms are mostly rejected into the residual glassy matrix.

Formation of the nanoscale icosahedral phase was observed in Zr–Cu–Al, Zr–Al–Ni–Cu [64], and Zr–Ti–Ni–Cu–Al [65] glassy alloys containing an impurity of oxygen above about 1800 mass ppm, although no icosahedral phase is formed if the oxygen content is lower than 1700 mass ppm. Recently the nanoscale icosahedral phase was obtained in Zr–Al–Ni–Cu–Ag [66], Zr–Al–Ni–Cu–Pd [67], Zr–Ni–NM [68] (NM-noble metals), Ti–Zr–Ni–Cu [69], Zr–Cu–Ti–Ni [70], Zr–Al–Ni–Cu–(V, Nb, or Ta) [71], Zr–Pd [72], and Zr–Pt [73, 74] system alloys at much lower (typically about 800 mass ppm) oxygen content. The icosahedral phase is also formed in $Zr_{70}Cu_{29}Pd_1$ alloy at very low Pd content [75]. It is also noted that the Zr-based amorphous alloys containing the NM elements exhibit the glass transition followed by the supercooled liquid region and hence the icosahedral phase precipitates from the supercooled liquid. At the same time, recently the nanoscale icosahedral phase has been produced in Zr–Al–Ni–Cu glassy alloys with low oxygen content below 500 mass ppm [76].

The study of the amorphous → icosahedral phase transformation in a $Zr_{65}Al_{7.5}Ni_{10}Cu_{7.5}Ag_{10}$ showed that the quasilattice constant decreases with annealing time and the

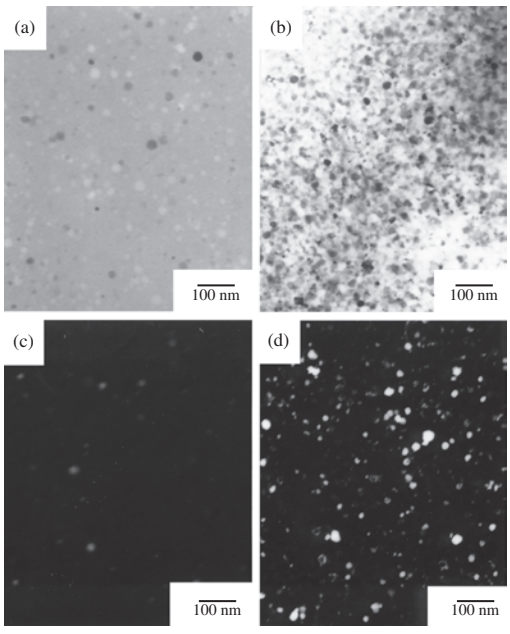


Figure 1. (a, b) Bright- and (c, d) dark-field TEM images. (a, c) After isothermal calorimetry at 725 K for 2200 s and (b, d) for 3500 s (scale marks are different). Reprinted with permission from [58], D. V. Louzguine and A. Inoue, *Scripta Mater.* 43, 371 (2000). © 2000, Elsevier Science.

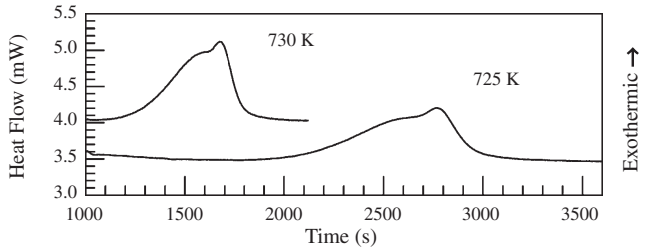


Figure 2. Isothermal differential calorimetry traces. Reprinted with permission from [58], D. V. Louzguine and A. Inoue, *Scripta Mater.* 43, 371 (2000). © 2000, Elsevier Science.

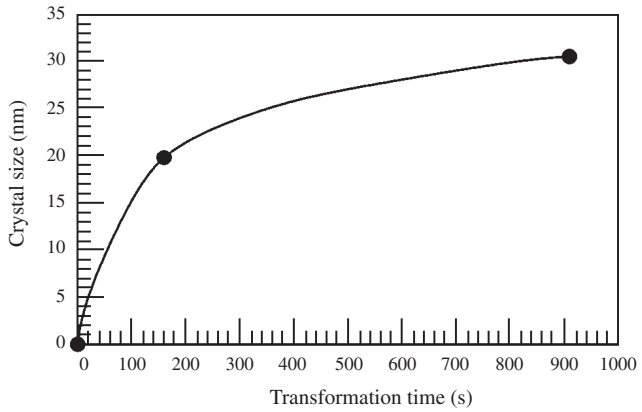


Figure 3. The largest Ti_2Ni particle size versus transformation time (without incubation period) plot.

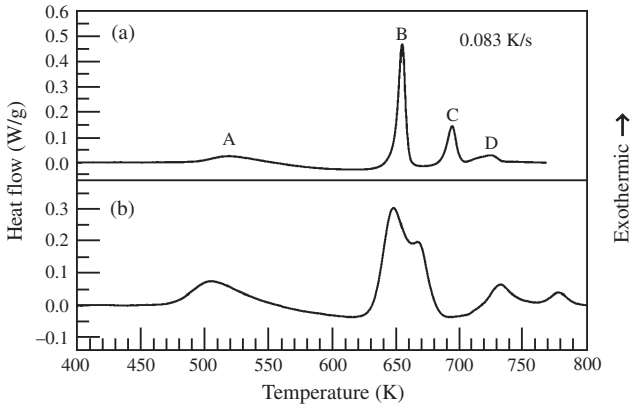


Figure 4. DSC traces of $Ge_{68}Cr_{14}Al_{10}Ce_4Sm_4$ and $Ge_{70}Cr_{16}Al_{10}Nd_4$ amorphous alloys taken at a heating rate of 0.083 K/s. Reprinted with permission from [80], D. V. Louzguine et al., *J. Mater. Sci.* 35, 5537 (2000). © 2000, Kluwer Academic Publishers.

zirconium and silver content in the icosahedral particles differs from those in the remaining amorphous matrix [77]. It confirms that this is a nonpolymorphous, rather primary-type reaction.

The local atomic structures around Pt as well as Zr in the amorphous and quasicrystalline $Zr_{70}Al_6Ni_{10}Pt_{14}$ alloys have been determined by the anomalous X-ray scattering method [78]. The existence of chemical short range order clusters was found around Pt atoms in the glassy state as well as in the quasicrystalline phase.

Two very important parameters of such composite nanomaterials are grain (particle) size and its volume fraction. Primary devitrification of highly supercooled liquid or amorphous phase often has the following unique features, that is, high nucleation frequency, low crystal growth rate, high concentration gradient of solute element at liquid/solid interface resulting from low atomic diffusivity, formation of metastable phases by the control of redistribution of solute elements, formation of a residual amorphous phase with high solute concentration, defect-free nanocrystalline particles with low residual strain, highly dense packed structure at liquid/solid interface, nanoscale interparticular spacing, and size and shape effects of nanoscale spherical particles. By utilizing these features, new nanoscale mixed-phase alloys containing a residual amorphous phase were obtained by devitrification of glassy alloys [79].

Various Ge–Cr–Al–RE amorphous alloys show a wide exothermic DSC peak preceding crystallization (Fig. 4). After the completion of this heat effect the structure remains almost amorphous, although medium range order zones below 1 nm in size were observed in heat treated as well as in the as-solidified state (Fig. 5, encircled).

The nature of the broad exothermic peak has been studied in the $Ge_{70}Cr_{16}Al_{10}Nd_4$ alloy by differential scanning calorimetry and X-ray diffraction [81]. The resultant coordination numbers for the first and second peaks of ordinary radial distribution function in the as-solidified state are 2.8 at 0.258 nm and 4.4 at 0.310 nm, respectively. The corresponding coordination numbers for the sample heat-treated up to the completion of the first broad heat effect are 3.3 at 0.250 nm and 2.9 at 0.308 nm, respectively [81]. The ratio between the second and the first nearest neighbor distances (0.41 and 0.250 nm, respectively) of 1.64 is almost identical with the value of 1.63 for *a*-Ge and *c*-Ge. The results indicate that during heat treatment tetrahedral configuration of Ge atoms which is close to the configuration observed in pure amorphous and crystalline Ge was formed. Formation of Ge-enriched zones precedes devitrification.

Finely dispersed nanoscale fcc Ge precipitates upon solidification of Ge–Al–TM–RE rapidly solidified alloys (Fig. 6) or upon devitrification of Ge–Al–TM–RE amorphous solids (Fig. 7) [82].

In Ge-based alloys one of the reasons for a nanoscale Ge particle size except for the diffusion-controlled growth discussed previously is elastic strain arising from volume changes upon precipitation. Tetrahedrally bonded crystalline Ge has higher specific volume compared to metallic liquid with a higher coordination number.

Precipitation of the nanoscale Ge–Si solid solution finely dispersed particles induced by Ge addition was observed in rapidly solidified $Si_{55-x}Ge_xAl_{20}Fe_{10}Ni_5Cr_5Zr_5$ alloys (Figs. 8 and 9) [83]. Average particle size of nanoscale Ge calculated

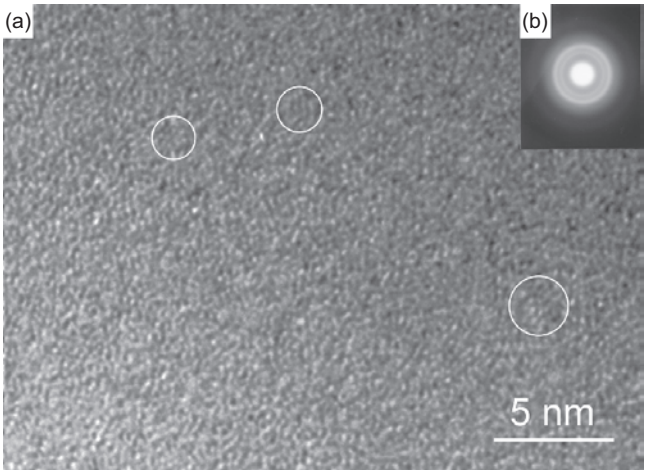


Figure 5. (a) High-resolution TEM image and (b) selected-area electron diffraction pattern of the $Ge_{68}Cr_{14}Al_{10}Ce_4Sm_4$ in the as-solidified state. Reprinted with permission from [80], D. V. Louzguine et al., *J. Mater. Sci.* 35, 5537 (2000). © 2000, Kluwer Academic Publishers.

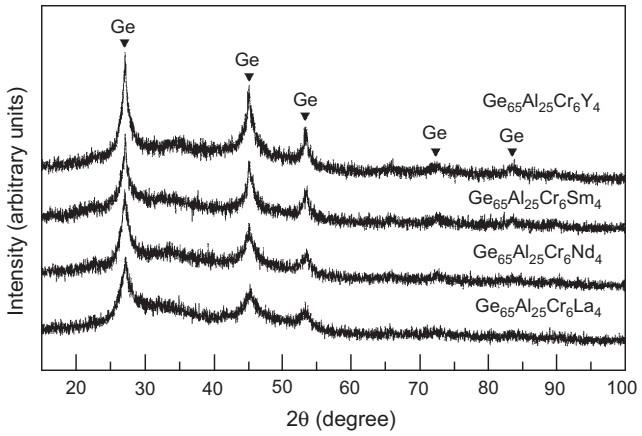


Figure 6. X-ray diffraction patterns of as-rapidly solidified $Ge_{65}Al_{25}Cr_6RE_4$ alloys. Reprinted with permission from [82], D. V. Louzguine and A. Inoue, *Mater. Trans. JIM* 40, 485 (1999). © 1999, Japan Institute of Metals.

from the broadening of the diffraction peaks increases with increase in Ge content and was found to be in good agreement with that obtained from TEM observations (Fig. 10). The smallest Ge particle size was obtained in $Si_{48}Al_{20}Fe_{10}Ge_7Ni_5Cr_5Zr_5$ alloy and these particles are finely dispersed (see Fig. 8).

The Ge diffraction peaks are shifted to higher 2θ angle values and lower interlattice distances d_{hkl} by the dissolved Si. The amount of Si dissolved in Ge decreases and the lattice parameter of Ge solid solution increases with increasing Ge concentration in the alloy as shown in Figure 11 [83]. Such a material consisting of nanosize particles of crystalline Ge solid solution (semiconductor with high electrical resistivity of about 4.7×10^5 $\mu\Omega$ m) in Si-based amorphous matrix has a resistivity of 11–13 $\mu\Omega$ m depending on the composition and can be possibly applied in electronic materials.

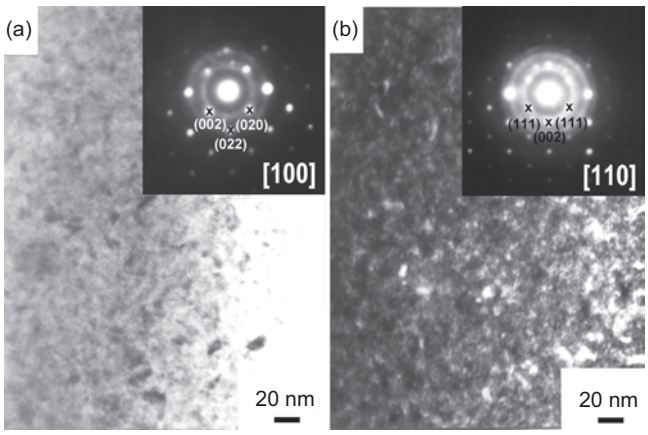


Figure 7. (a) Bright-field and (b) dark-field TEM images of $Ge_{55}Al_{30}Cr_{10}La_5$ alloy heat treated for 0.3 ks at 595 K. The insets are selected area electron diffraction patterns taken from the largest Ge particles, their indexing and zone axis. Reprinted with permission from [82], D. V. Louzguine and A. Inoue, *Mater. Trans. JIM* 40, 485 (1999). © 1999, Japan Institute of Metals.

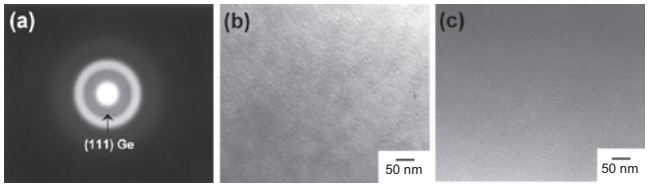


Figure 8. (a) Selected area electron diffraction pattern, (b) bright-field and (c) dark-field electron micrographs of rapidly solidified $Si_{48}Al_{20}Ge_7Fe_{10}Ni_5Cr_5Zr_5$ alloy. Reprinted with permission from [83], D. V. Louzguine and A. Inoue, *Mater. Trans. JIM* 39, 245 (1998). © 1998, Japan Institute of Metals.

Structure types that can be obtained in rapidly solidified Si, Ge, Al, and TM containing alloys are summarized in Table 1 [84].

3. CLASSIFICATION

As the first approximation dimensionality of nanomaterials (in how many of three dimensions a material has a nanoscale size) can be used for their classification. Thin layers, multilayers, coatings, and quantum well structures can be treated as one-dimensional nanomaterial; thin wires [85, 86] or nanotubes (carbon [87] or BN [88], for example) as two-dimensional nanomaterial; and submicrometer grains, isolated particles, fullerenes [89], and quantum dots as three-dimensional nanomaterials, that is, they have a nanoscale size in any of the three dimensions (atomic clusters can also fit into this group). However, nanoscale coatings and films on the surface of bulk materials being actually one-dimensional can be considered as a special class of nanomaterial. Bulk polycrystalline nanostructured materials having a size exceeding 1 mm in any dimension nevertheless consist of nanoscale grains. Nanocomposites [90] are another quickly developing branch in the field of nanomaterials. A classical example is age-hardened Al-based alloys consisting of an aluminum solid solution matrix reinforced with dispersed (tens of nanometers in size) nanoparticles.

H. Gleiter, one of the founders of the nanomaterials field, proposed another classification scheme consisting of three categories [91]. The first category contains materials or devices with reduced dimensions in the form of (isolated, substrate-supported, or embedded) nanometer-sized particles, thin wires, or thin films. The second category comprises materials in which the nanometer-sized microstructure is

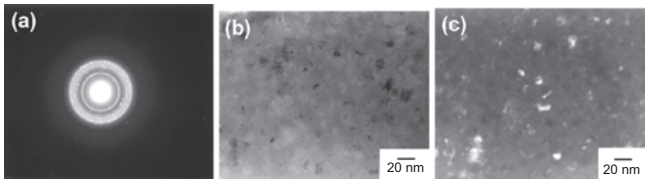


Figure 9. (a) Selected area electron diffraction pattern, (b) bright-field and (c) dark-field electron micrographs of rapidly solidified $Si_{45}Al_{20}Ge_{10}Fe_{10}Ni_5Cr_5Zr_5$ alloy. Reprinted with permission from [83], D. V. Louzguine and A. Inoue, *Mater. Trans. JIM* 39, 245 (1998). © 1998, Japan Institute of Metals.

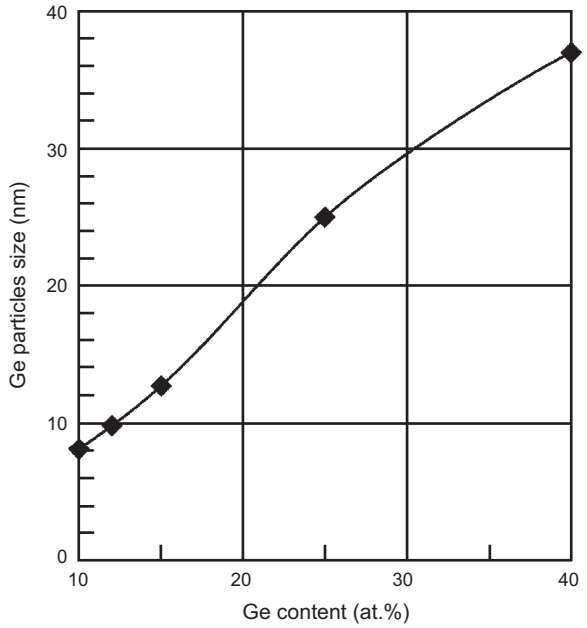


Figure 10. Average particle size of *c*-Ge as a function of Ge content for $Si_{55-x}Al_{20}Ge_xFe_{10}Ni_5Cr_5Zr_5$ alloys. Reprinted with permission from [83], D. V. Louzguine and A. Inoue, *Mater. Trans. JIM* 39, 245 (1998). © 1998, Japan Institute of Metals.

limited to a thin (nanometer-sized) surface region of a bulk material. Third category includes bulk solids with a nanoscale microstructure grain, crystallite, dendritic cell size, etc. There are two classes of such solids. In the first class, the atomic structure and/or the chemical composition varies in space continuously throughout the solid on an atomic scale. Glasses, gels, supersaturated solid solutions, or implanted

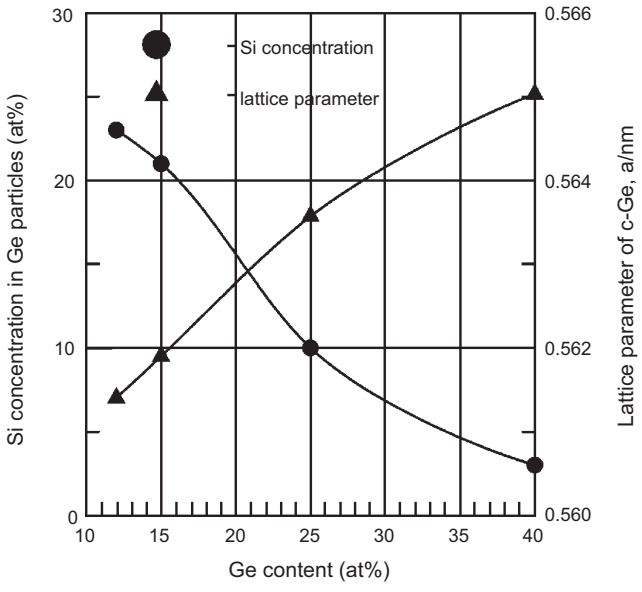


Figure 11. Si concentration in *c*-Ge and the lattice parameter of Ge solid solution as a function of Ge concentration for $Si_{43-15}Ge_{12-40}$. $Al_{20}Fe_{10}Ni_5Cr_5Zr_5$ alloys. Reprinted with permission from [83], D. V. Louzguine and A. Inoue, *Mater. Trans. JIM* 39, 245 (1998). © 1998, Japan Institute of Metals.

Table 1. Structure types that can be obtained in rapidly solidified Si, Ge, Al, and TM containing alloys according to Si and Ge content.

	System	
Structure	Si–Al–TM and Si–Al–TM–Ge	Ge–Al–TM and Ge–Al–TM–Si
Amorphous	50–55% Si[a] (<5% Ge)	50% Ge (<2% Si)
Amorphous + nanocrystalline Ge (<20 nm)	40–55% Si, 7–15% Ge	55% Ge
Amorphous + crystalline Ge (>20–30 nm)	30–5% Si, 25–50% Ge	47–45% Ge, 3–5% Si

[a] Atomic percent.

Source: Reprinted with permission from [84], D. V. Louzguine and A. Inoue, in "Amorphous and Nanocrystalline Materials, Preparation, Properties, and Application" (A. Inoue and K. Hashimoto, Eds.), Advances in Materials Research, Vol. 3. Springer-Verlag, Berlin/Heidelberg, 2001. © 2001, Springer-Verlag.

materials are examples of this type. A second class of materials with a nanoscale microstructure are assembled of nanoscale size structure units, mostly crystallites. These structure units may differ in their atomic structure, their crystallographic orientation, and/or their chemical composition. If the units are crystallites, incoherent or coherent interfaces may be formed between them, depending on the atomic structure, the crystallographic orientation, and/or the chemical composition of adjacent crystallites. In other words, materials assembled of nanometer sized structure units are microstructurally heterogeneous consisting of the grains (crystallites) and the regions between adjacent grains (e.g., grain boundaries). This inherently heterogeneous structure on a nanometer scale is crucial for many of their properties and distinguishes them from glasses, gels, etc. that are microstructurally homogeneous. One should also say that glassy alloys phase separated by a spinodal mechanism represent a nanomaterial with two kinds of glassy phases.

Although nanostructured materials are mostly in nonequilibrium state equilibrium nanomaterials also exist as self-assembled structures. Supramolecules are produced by the intermolecular association of receptors and substrates following an inherent assembling pattern based on the principles of molecular recognition. Supramolecular self-assembly concerns the spontaneous association of a number of components resulting in the generation of supermolecules or extended polymolecular assemblies such as molecular layers and films [92, 93]. However, in this chapter we will focus on nanostructured metals and alloys only.

4. STRUCTURE

Knowledge of the structure of nanomaterials, especially their grain boundaries, is critical for understanding their properties. Among the techniques used for structure characterization of nanomaterials are X-ray diffractometry (XRD), TEM, including (HRTEM) imaging, Mössbauer spectrometry, etc. [94]. HRTEM is a direct and extremely powerful method for observation of nanostructure. The atom probe technique also has been used for materials characterization and demonstrated a unique capability of characterizing local chemical compositions of nanoscale metallic microstructures. In particular, the three-dimensional atom

probe (3DAP) technique gives three-dimensional mapping of all alloying elements in a real space with a near atomic resolution. Thus, HRTEM and 3DAP are the most suitable techniques for analyzing composition of nanoscale precipitates, zones, or clusters. Devitrification in Zr-based bulk metallic glasses has been directly detected using high intensity, high energy monochromatic synchrotron beam diffraction in transmission during *in-situ* heat treatment [95].

4.1. Nanocrystalline Alloys

An important problem is precise determination of the nanocrystal size. TEM allows direct imaging and calculation of the grain size and is one of the most powerful techniques for structural investigation of nanomaterials. Moreover, nanobeam electron diffraction and energy dispersive X-ray spectrometry (EDX) allow direct analysis of the crystal structure and chemical composition of nanoparticles. However, the grain overlap problem sometimes produces difficulties in determination of the average grain size. Analysis of the broadening of the diffraction peaks obtained by XRD gives an average crystal size of the whole sample [96]. However, broadening of the XRD diffraction peaks is caused by both small grain size and microdeformation. Although there are several methods to analyze and separate an income of the grain size and microdeformation to the broadening of the diffraction peak [97–99] the resulting grain size differs from method to method [100].

X-ray-diffraction measurements performed on a nanocrystalline Cu sample made by severe plastic deformation showed that the shape of the diffraction peaks was found to be represented primarily by a Lorentzian function. A small difference was revealed between the integrated intensities from the nanocrystalline and a reference Cu sample. The broadening of the diffraction peaks of the nanocrystalline Cu sample was induced by both small crystallite size and microstrain inside the grains [101]. Although at the grain size below 50–100 nm dislocations easily annihilate on the grain boundaries the lattice of nanomaterials obtained by severe plastic defloration is heavily distorted elastically. It was found that the grain sizes of nanocrystalline Cu in different crystallographic orientations are essentially the same, while the microstrain exhibits a significant anisotropy. Nanomaterials obtained by severe plastic deformation show a high level of elastic strain in the grains and in the areas close to grain boundaries the elastic strain level is much higher than in the central regions.

Due to a very small grain (structure unit) size nanostructured materials have a high fraction of grain boundaries. The volume fraction of atoms directly influenced by grain boundaries can be estimated by the formula $V_g = 3e/\langle d \rangle$, where e is an average effective thickness of the zone of influence of grain boundaries and $\langle d \rangle$ is the average grain size [102]. For $e = 1$ nm and $\langle d \rangle = 10$ nm V_g is $\sim$30%. High volume fraction of atoms at grain boundaries gives a large contribution to Mössbauer spectra. Although Mössbauer spectrometry is a less direct technique than HRTEM it provides complementary information on statistical features which are difficult to obtain using TEM [103]. It has been also shown that the coordination number of nanostructured Pd as a function of the interatomic spacing is different from that of a conventional polycrystalline Pd [104] which is also connected with the high volume fraction of grain boundaries.

The grain boundaries in nanomaterials have been also studied by TEM. Initially it was considered that grain boundaries in nanomaterials are thicker than those of the conventional alloys [91]. However, high resolution TEM studies (Fig. 12) showed that thickness of grain boundaries, for example in nanostructured Pd [105], does not exceed 1–2 interatomic distances.

An important question is do grain boundaries in nanomaterials differ from that in conventional materials? As reported in [106, 107] grain boundaries in nanomaterials have the short range order of grain boundaries formed in conventional polycrystalline materials. However, other studies showed that the grain boundaries in nanomaterials, especially obtained by plastic deformation, contain defects and thus have high energy and nonequilibrium character [108, 109]. One can say that the type of grain boundary (equilibrium or nonequilibrium [110]) also depends upon the production method of the nanomaterial.

All the aforementioned also supports the idea that nanostructured materials are different from conventional polycrystalline materials.

4.2. Nano-Quasi-Crystalline Alloys

An icosahedral phase is formed in rapidly solidified Al–TM based alloys containing Mn [111], Cr [112], and V [113, 114] as TM elements. It is possible to obtain nanostructure by using a high cooling rate or severe plastic deformation. The Al–TM base icosahedral structure has been presumed [115] to consist of Mackay icosahedral clusters containing 54 atoms which are arranged through glue atoms into a three-dimensional quasiperiodic structure. Nanomaterial consisting of icosahedral almost spherical particles with diameters below 10 nm was obtained in Zr–Pt alloy [26] by casting. The formation of the nanoscale icosahedral phase during casting indicates that the icosahedral short-range order exists in the melt of the Zr–Pt binary alloy. For example, the structure of the $Zr_{65}Al_{7.5}Ni_{10}Cu_{12.5}Pd_5$ alloy annealed for 18 s at 725 K is shown in Figure 13.

Local atomic structures around NM as well as Zr in amorphous and quasicrystalline Zr-based alloys have been

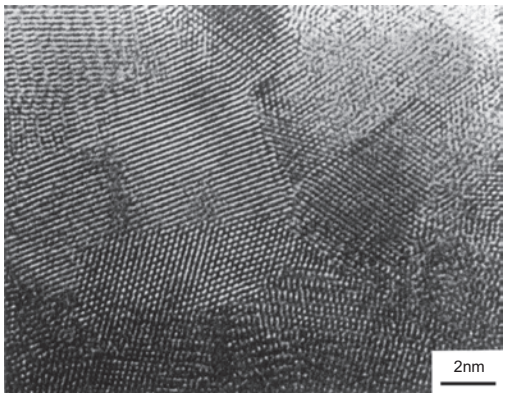
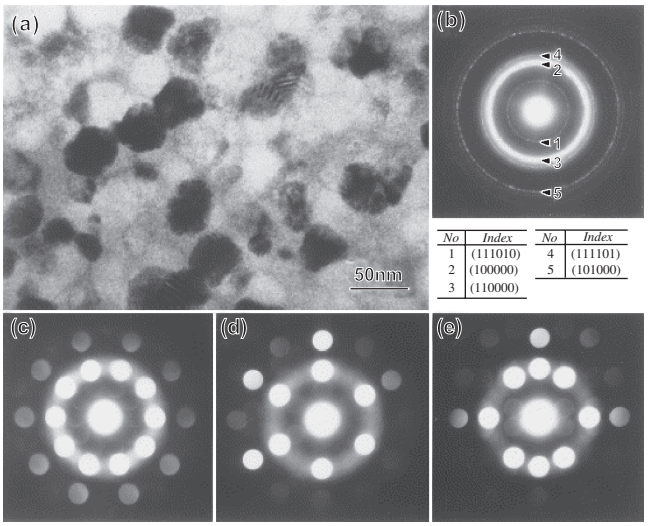


Figure 12. HRTEM image of nanocrystalline Pd. Reprinted with permission from [105], G. J. Thomas et al., *Scr. Metall. Mater.* 24, 201 (1990). © 1990, Elsevier Science.



No	Index	No	Index
1	(111010)	4	(111101)
2	(100000)	5	(101000)
3	(110000)		

Figure 13. Bright-field TEM image (a), selected-area electron diffraction pattern (b), and nanobeam diffraction patterns (c)–(e) of the melt-spun glassy $Zr_{65}Al_{7.5}Ni_{10}Cu_{12.5}Pd_5$ alloy annealed for 18 s at 725 K. Reprinted with permission from [116], A. Inoue et al., *Scripta Mater.* 44, 1615 (2001). © 2001, Elsevier Science.

determined by anomalous X-ray scattering. The existence of chemical short range order clusters that are formed around the NM atom in the amorphous as well as in the quasicrystalline phase [78] is confirmed.

Quasicrystals can often coexist with rational approximant crystalline structures. A multiple twinning around an irrational axis of the approximants has been reported in an aggregate of fine size cubic crystallites. The crystalline aggregate symmetry is linked to the icosahedral point group symmetry ($m\bar{3}\bar{5}$) [117].

4.3. Nanostructured Alloys with Mixed Glassy and (Quasi)Crystalline Structure; Nanocomposites

Nanoscale particles of a crystalline or quasicrystalline phase can be easily produced by devitrification of glassy alloys. Figure 14 summarizes the features of the microstructure of nonequilibrium Al-based alloys. The nonequilibrium structures for Al-based alloys are classified into the following six types, (1) amorphous single phase, (2) nanostructure consisting of Al and compounds obtained by devitrification of amorphous phase, (3) partially crystallized structure of nanoscale fcc-Al particles embedded in an amorphous matrix, (4) nanoquasicrystalline structure consisting of nanoscale quasicrystalline particles surrounded by Al phase without grain boundary, (5) coexistent nanogranular amorphous and Al phases, and (6) nanogranular Al phase surrounded by an amorphous network.

The structure types of Zr-based nanostructured alloys with mixed glassy and (quasi)crystalline structure are shown in Figure 15.

One should also mention that some bulk glassy alloys, for example $Zr_{55}Co_{20}Al_{12}Cu_5Nb_8$, contain medium-range order (MRO) zones and nanoscale particles in as-solidified state (Fig. 16) even though these precipitates do not produce

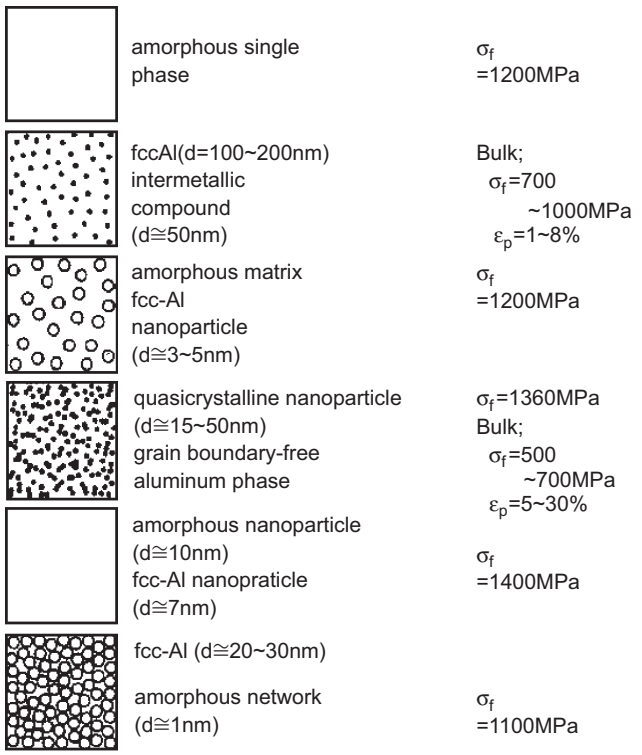


Figure 14. Microstructure and mechanical properties of nonequilibrium Al-based alloys. Reprinted with permission from [118], A. Inoue and H. M. Kimura, *J. Metastable Nanocryst. Mater.* 41 (2001). © 2001, Trans Tech.

diffraction peaks in the XRD and selected-area electron diffraction pattern due to their low volume fraction.

4.4. Grain Growth and Other Structure Changes on Heating Except for Phase Transformations

The driving force of grain growth in metal or alloy is the excess of free energy. Although nanoscale grains in nanostructured metals and alloys usually contain no dislocations they have an extremely large fraction of atoms in grain boundaries. In nanomaterials obtained by severe plastic deformation these grain boundaries have a nonequilibrium nature and have grain boundary dislocations [121]. These dislocations cause a high level of elastic strain.

Grain growth in nanostructured materials obtained by severe plastic deformation having highly nonequilibrium grain boundaries starts at quite a low reduced temperature [122, 123].

Grain growth in nanostructured metals and alloys has been studied [124, 125]. Structure changes in nanocrystalline Fe on heating were studied by X-ray diffraction and DSC [126] and it is found that the n exponent in the equation [127, 128]

$$D^{1/n} - D_0^{1/n} = k't \tag{2}$$

for a mean grain diameter D after time t where k' is the rate constant deviates from (0.5) what it should be in the

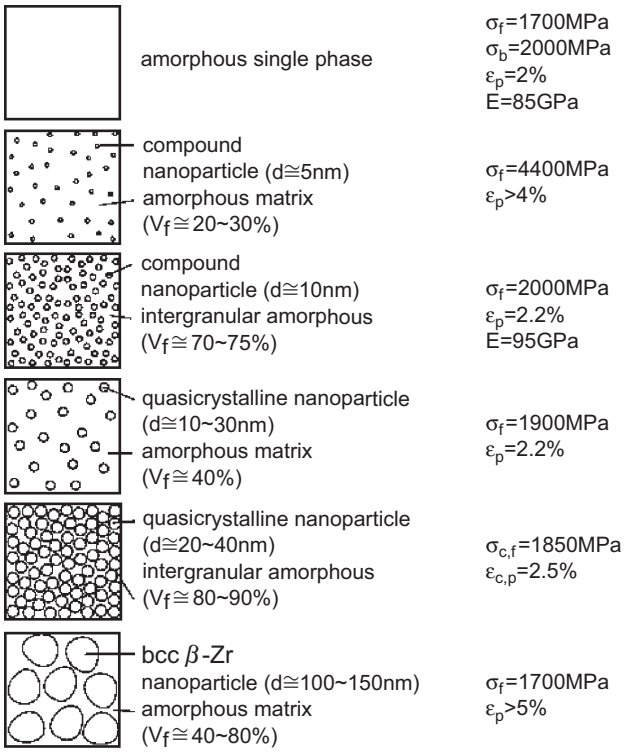


Figure 15. Microstructure and mechanical properties of nonequilibrium Zr-based bulk alloys in Zr–Al–Ni–Cu and Zr–Al–Ni–Cu–M (M = Ag, Pd, Au, Pt, or Nb) systems. Reprinted with permission from [119], A. Inoue et al., *Sci. Technol. Adv. Mater.* 1, 73 (2000). © 2000, Elsevier Science.

case of parabolic growth. The same results were obtained by other researchers [129, 130].

Another process taking place on heating is the relaxation process leading to reduction of dislocation density. For example, this process took place under heating of nanostructured Ni obtained by severe plastic deformation and was followed by recrystallization [131].

Nanoparticles obtained by devitrification of glassy phase in many cases show quite high thermal stability against coarsening [132]. The grain growth curve of cF96 Ti_2Ni crystalline nanoparticles in Ti-based glassy matrix shows significant retardation of the particle growth rate when they reach a size of about 20 nm (Fig. 3) due to diffusion controlled growth. At the diffusion controlled growth a particle radius r or diameter D is proportional to $t^{1/2}$. Fitting of the particle size versus time curve (Fig. 3) gives an exponent value of 0.25. Such a low value is possibly due to rejection of Cu from the Ti_2Ni phase forming a Cu-enriched layer in the Ti_2Ni/residual glassy matrix interface. The residual glassy matrix further forms γTiCu phase.

Inclusions of nanoparticles in large grains as well as segregation of some atoms on the grain boundaries hampers grain growth. Grain growth of pure Ni and Ni–Si solid solution was studied by DSC [133]. It was found that the segregation of Si atoms on the grain boundaries impedes growth of the nanocrystals. It has been also shown that [134] nanoscale grains may reach a metastable state where the structure is stable with respect to variation of its total grain boundary area and the grain boundary free energy is significantly reduced.

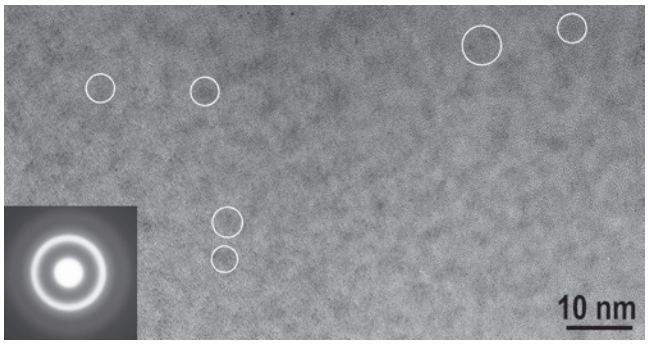


Figure 16. HRTEM image of rod-shaped bulk (2 mm in diameter) $Zr_{55}Co_{20}Al_{12}Cu_5Nb_8$ glassy alloy containing MRO zones and nanoscale particles in as-solidified state. The insert represents selected-area electron diffraction pattern. Reprinted with permission from [120], T. Wada, Master's Thesis, Tohoku University, 2003.

5. PROPERTIES

Properties of nanomaterials are quite different from those of conventional crystalline alloys as well as from the amorphous and glassy solids. For example, due to high fraction of grain boundaries nanocrystalline materials have lowered density compared to that of conventional polycrystalline materials. It has been shown that Curie temperature and other magnetic properties of nanostructured Ni also change when a grain size becomes nanoscale [135]. Not only are mechanical strength and ductility of Cu strongly affected when grain size reaches nanoscale values but also such a non-structure-sensitive properly as Young's modulus [136]. The aforementioned supports an idea of separating nanomaterials into a special class of solid substances. In the present section we will consider mechanical and magnetic properties of metallic nanomaterials.

5.1. Mechanical Properties

Nanostructured metals and alloys exhibit mechanical properties which are significantly different from the properties of conventional polycrystalline alloys. They possess high hardness, tensile, and compressive strength values and exhibit superplastic behavior at low reduced temperature.

5.1.1. General Description of Deformation Behavior on the Basis of Crystalline Single-Phase Nanostructured Alloys

A number of works are dedicated to theoretical and experimental understanding of mechanical behavior of nanostructured metals and alloys. In general bulk nanostructured materials were found to exhibit low ductility at room temperature that imposes a limit on their practical applications [137]. For example, pure copper which is naturally very ductile became quite brittle when its grain size decreased below 50–100 nm. Especially low plastic elongation values of 1–2% are obtained for materials with a grain size below

25 nm. For example, electrodeposited nanocrystalline Ni with a mean grain size of 10 nm showed yield strength, ultimate tensile strength, and elongation to fracture of 900 MPa, 2000 MPa, and 1%, respectively compared to 100 MPa, 400 MPa, and 50%, respectively, for polycrystalline Ni with 10 μm grain size [138, 139]. The major reasons for brittleness of bulk metallic nanomaterials are connected with the porosity, weak bonding of the grains/particles and other artifacts, as well as instability of crack nucleation and propagation. Limited ductility and toughness of nanomaterials is closely connected to the absence of any significant work hardening. However, recent studies on nanocrystalline pure metals showed that quite high ductility values on tensile test can be obtained at a definite grain size after proper treatment [140]. For example, nanocrystalline hcp Co samples with a mean grain size of 12 nm produced by pulse electroplating on Ti substrate followed by a thermomechanical treatment showed an ultimate tensile strength of 1865 MPa and yield strength of 1000 MPa at 7% elongation which is an example of a very good combination of strength and ductility [141].

Mechanical yield strength of conventional polycrystalline materials obeys the well-known Hall–Petch relationship that has been extrapolated to nanomaterials as well [142]. The Hall–Petch relation states that the yield strength σ_y increases monotonically with decreasing the average grain size d as

$$\sigma_y = \sigma_0 + kd^{-1/2} \tag{3}$$

where σ_0 is the friction stress needed to move dislocations and k is a material dependent constant. Considerable strengthening is achieved on reducing the grain size from the macro- to the micro- and finally to the nanoscale size. However, different types of strength variation with nanoscale grain size have been reported, ranging from respecting the Hall–Petch relationship to showing a saturation of strength or even fall of strength at the finest grain sizes. Several Hall–Petch plots of yield stress or hardness against the reciprocal of the square root of the average grain size show a deviation from straight line behavior when grain size is lower than 25 nm [143–145].

A number of models have been proposed to interpret the inverse Hall–Petch behavior. Some calculations [146] based on the idea that very small grains cannot sustain significant dislocation activity indicate that Hall–Petch behavior should cease when the average grain size is below 10–20 nm. Numerous models have been devised and molecular dynamics simulations have been carried out to explain deformation when the grain size falls below the cutoff value for significant dislocation activity. It is frequently predicted by a mechanism based on grain boundary sliding accommodated by diffusion or by the emission of dislocations into the grains [147, 148]. However, recent *in-situ* observations of deformation behavior in nanocrystalline Cu by TEM have shown dislocation motion mechanism of plastic deformation [149]. No evidence for grain boundary sliding or rotation in the range of grain sizes that could be studied in the present experiment was observed. The inverse Hall–Petch relation is also explained by the size dependent line tension where the dislocation elastic strain is assumed to be screened within the distance of the grain diameter [150]. The existence of a Frank–Read source in large grains in nanocrystals was also assumed [151]. On the basis of the composite model involving a crystalline matrix and intercrystalline layers, the yield stress of nanocrystalline materials can be expressed as the sum of the matrix contribution and the intercrystalline contribution, both of which are assumed to be linearly connected with the corresponding shear module and to be functions of the grain size [152]. An attempt was also done to explain the inverse Hall–Petch relation by the change of the deformation mechanism from a dislocation glide process to the Coble creep [153]. In this model the shape of the yield stress against grain size curve depends on the relative values of Hall–Petch slope, the conventional Coble constant, the threshold constant, and the width of the grain size distribution.

Creep behavior of nanocrystalline Cu and Pd has been studied over a range of temperatures (0.24–0.64 T_m) on the samples prepared by inert gas condensation and compaction [154]. The measured creep rates were found to be two to four orders of magnitude smaller than the values predicted by the equation for Coble creep. At the same time a low temperature tensile creep test of nanoscale Cu with an average grain size of 30 nm prepared by electrodeposition and investigated in the temperature range of 0.22–0.24 T_m showed that the creep rates are of the same order of magnitude as those calculated from the equation for Coble creep [155]. The results suggest that the creep can be attributed to the "interface controlled Coble creep." The difference in creep behavior between electrodeposited Cu and that produced by inert gas condensation can be attributed to different preparation techniques. Once again one can point out that the preparation method has an influence on the properties of the product.

Fatigue properties of nanocrystalline Cu prepared by inert gas condensation [156] (an increase of grain size by 30% is observed after cycling) and severe plastic deformation [157] have been also studied. No cycling softening and high saturation stress were observed. However, nano-submicrometer-scale Cu produced by severe plastic deformation in [158] tested in tension–compression fatigue under strain control showed significant softening during testing.

5.1.2. Crystalline Multiphase (Two and More) Nanostructured Alloys

Commercial materials usually have a multiphase microstructure.

When the extrusion temperature is raised to the second-stage exothermic reaction range, one can easily produce a fully dense nanocrystalline alloy consisting of fine compounds with a size of about 50 nm embedded in an Al matrix with a grain size of 100 to 200 nm [159]. The nanocrystalline alloys exhibit a good combination of high tensile strength above 900 MPa and high fatigue strength of 300 to 350 MPa which are much superior to those for conventional Al-based alloys as well as for new Al-based alloys [160] developed by rapid solidification and powder metallurgy techniques. Figure 17 shows the relation between the yield strength ($\sigma_{0.2}$) and the grain size of the fcc-Al matrix phase for the

nanocrystalline Al-based alloys, together with the data of conventional crystalline Al-based alloys. The present high tensile strength has been interpreted to result from the combination of two mechanisms of dispersion strengthening and grain size refinement strengthening. The combination of the two strengthening mechanisms is attributable to the formation of the unique mixed structure in which intermetallic compounds disperse homogeneously within the Al grains and on the grain boundaries. The dispersion density of the compounds is not high enough to suppress all the dislocations and a significant part of dislocations is hindered at grain boundaries of Al phase. Figure 17 also shows that the increase in the yield strength by the dispersion strengthening and grain size refinement strengthening is approximately 400 and 390 MPa, respectively. When the dispersion strengthening is attributable to the Orowan mechanism [161], the increase in the yield strength is evaluated by

$$\Delta\tau = 0.81Gb/\{2\pi(1-\nu)^{1/2}\} * \ln(2r_s/r_0)/(\lambda_s - 2r_s) \quad (4)$$

Here, G is shear modulus of Al (27 GPa), b is Burger's vector of Al (2.86×10^{-10} m), ν is Poisson's ratio of Al (0.345), r_s is particle radius $\{2/3(r)\}$, r_0 is core dislocation radius for Al (2.86×10^{-10} m), λ_s is $1.25r(2\pi/3f)^{1/2}$, and f is volume fraction of the dispersoids. The r and f for the present alloy are measured to be 2.5×10^{-8} m and 0.34 [162], respectively. The $\Delta\tau$ is calculated to be 129 MPa and the resulting increase of yield stress $\Delta\sigma$ is 395 MPa.

Furthermore, the increase in σ_y by the grain size refinement strengthening is expressed by the Hall–Petch relation (3). Here, σ_0 is yield strength of a single grain due to all strengthening mechanisms except grain boundary contributions [163], k is a constant relating to how effective the grain boundaries are in increasing yield strength, d is grain or subgrain size, and n is an exponent in the range 0.5 to 1 and is assumed to be 0.5 [162]. The σ_0 is due to solid solution strengthening and strain strengthening up to 0.2% elongation and included into the dispersion strengthening. The k value has been reported to be 75 MPa $\mu m^{1/2}$ for commercial pure Al [164], 120 MPa $\mu m^{1/2}$ for optimally aged 7075 [165], and 220 MPa $\mu m^{1/2}$ for underaged 7091 [166]. In the calculation, the k value was assumed to be 135 MPa $\mu m^{1/2}$. The resulting kd^{-n} value is 426 MPa for $d = 100$ nm and 302 MPa for $d = 200$ nm. Consequently, the yield strength ($\sigma_{0.2}$) evaluated by the summation of the dispersion strengthening and grain size refinement strengthening is 821 MPa for $d = 100$ nm and 697 MPa for $d = 200$ nm. These calculated $\sigma_{0.2}$ values agree with the experimental results shown in Figure 17, indicating an appropriateness of the strengthening mechanisms.

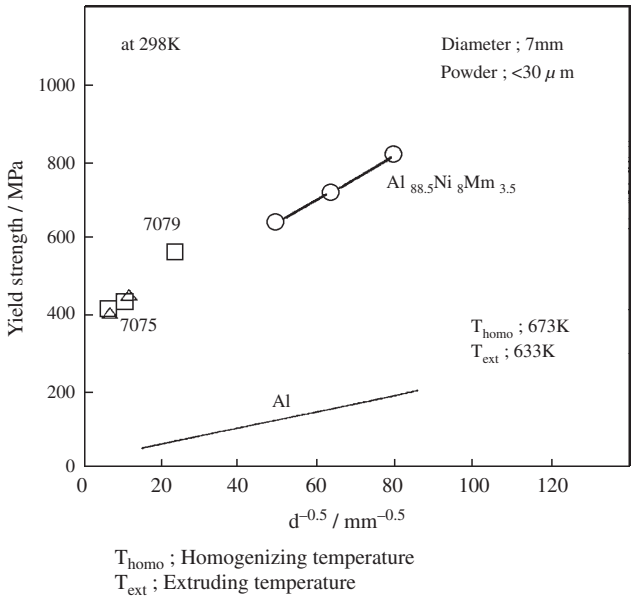


Figure 17. Relation between the yield strength ($\sigma_{0.2}$) and the grain size of Al matrix phase ($d^{-1/2}$) for nanocrystalline Al–Ni–Mm base alloys prepared by the powder metallurgy technique. The data of conventional Al-based alloys are also shown for comparison. Reprinted with permission from [167], A. Inoue and H. Kimura, *J. Light Metals* 1, 31 (2001). © 2001, Elsevier Science.

Mg-based solid solutions with high Mg content above 90 at% were produced by warm extrusion of gas atomized powders [rapid solidification powder metallurgy (RS P/M)] [168]. The structure of these high-strength RS P/M Mg-based alloys consisted of hcp solid solution containing fine compounds [169]. However, the resulting mixed phase alloys exhibit low elongation less than 3% and the poor ductility has prevented the progress of the Mg-based alloys as structural materials.

The compositional dependences of microstructure, Vickers hardness, and bend ductility for melt-spun Mg–Ln–TM and Mg–Y–Mm alloys were examined with the aim of finding an alloy composition with higher hardness at a good ductility. By choosing alloys with good bend ductility and higher hardness number consolidated bulk alloys were produced by warm extrusion of high pressure gas atomized alloy powders, that is, RS P/M process. The process consisted of the production of alloy powders by high-pressure gas atomization, followed by collection, sieving, precompaction, sealing into a Al can, and then extrusion at an extrusion ratio of 10 in the temperature range of 573 to 723 K [169].

The tensile yield strength at room temperature for the bulk $Mg_{95.5}Mm_2Y_{2.5}$ alloys produced at different extrusion temperatures ranging from 573 to 698 K exceeds 500 MPa in the extrusion temperature range of 573 to 648 K, but the elongation is less than 1.0% [170]. The consolidated bulk alloys of $Mg_{95.5}Y_2Mm_{2.5}$ alloy made from the atomized powders have two kinds of structure: (1) hcp solid solution single phase and (2) a mixed structure consisting of fine compounds embedded in hcp-Mg matrix phase. The low elongation is a disadvantage for the bulk hcp Mg-based alloys in Mg–Mm–Y and Mg–Ln–TM systems. On the contrary, the Mg-based solid solution alloys exhibit very high elevated temperature strength. For $Mg_{95}Mm_3Zn_2$ and $Mg_{95.5}Mm_2Y_{2.5}$ alloys the high yield strength above 400 MPa is kept in the temperature range up to 523 K and the strength level is 2 to 3 times higher than that [171] for the commercial WE54A-T6 alloy in a wide temperature range below 550 K. A similar high elevated temperature strength was also obtained for extruded bulk $Mg_{95}Ca_{2.5}Zn_{2.5}$ and $Mg_{96.5}Ca_1Zn_{2.5}$ alloys, though their elongation is also lower than 3% [172].

Based on the previous knowledge that the low elongations below 3% prevent practical uses of the Mg-based

alloys, a completely new Mg-based alloy with much higher $\sigma_{0.2}$ and larger δ values was developed [173]. In the development, the attention was paid to new ternary Mg-based systems which satisfy the following three criteria: (1) a multicomponent system consisting of more than three elements, (2) significant atomic size ratios above about 12% among the three elements, and (3) suitable negative heats of mixing among their elements. It has generally been approved that the three criteria are very effective for the stabilization of metallic supercooled liquid [174]. The effect of the stabilization is expected to cause the formation of a solid solution with a higher degree of nonequilibrium state. Mg–Al–Zn, Mg–Ca–Al, and Mg–Zn–Y systems satisfy the three criteria. The composition dependence of H_v and bend ductility were examined systematically in the annealed states for the melt-spun Mg–Ca–Al and Mg–Zn–Y alloys. The higher H_v value at the maintenance of good bend ductility was obtained for $Mg_{85}Ca_5Al_{10}$ and $Mg_{97}Zn_1Y_2$ alloys.

By choosing these two alloys, alloy powders with an average size of about 25 μm were produced by high pressure gas atomization. The atomized powders were subsequently collected, sieved, precompacted, sealed into a Al can, and then extruded at an extrusion ratio of 10 in the extrusion temperature range of 573 to 723 K.

The atomized $Mg_{85}Ca_5Al_{10}$ powder with a size below 25 μm consisted of hcp-Mg, Al_2Ca, and unknown nonequilibrium phases. The structure changed to hcp-Mg and Al_2Ca phases for the bulk alloy obtained by extrusion at 623 K. The extruded bulk alloy exhibited rather good mechanical properties. The tensile yield strength varied from 420 to 500 MPa at 5 to 8% elongation. It is noticed that the ductility value increases significantly without detriment to the high strength. The bulk Mg–Ca–Al alloy has a mixed structure of fine hcp–Mg grains with a size of about 400 nm including much smaller precipitates with an average size of about 50 nm.

A similar experiment was made for the $Mg_{97}Zn_1Y_2$ and $Mg_{96}Zn_1Y_3$ alloys [175]. The atomized $Mg_{97}Zn_1Y_2$ powder with an average size below 32 μm had a mixed structure consisting mainly of hcp-Mg and $Mg_{24}Y_5$ phases. The Mg–Zn–Y alloy powder can be easily consolidated into a bulk form with high density. The density is 1.84 g/cm^3 for the 2% Y alloy and 1.86 g/cm^3 for the 3% Y alloy. Figure 18 shows the $\sigma_{0.2}$ and δ values as a function of extrusion temperature for the RS P/M $Mg_{97}Zn_1Y_2$ alloy. As the extrusion temperature rises, the $\sigma_{0.2}$ decreases from 600 to 420 MPa while the δ increases from 5% to 15%. It is noticed that the $\sigma_{0.2}$ and δ values are about 1.5 times higher than those for the former Mg–Y–Mm and Mg–Zn–Ln alloys. The specific yield strength values of the present Mg–Zn–Y and Mg–Ca–Al alloys are in the range of 2.7×10^5 to 3.3×10^5 Nm/kg which is about 1.2 to 4 times higher than those for the conventional Mg, Al, and Ti alloys [176].

One can note that one of the national projects is the development of Al-based amorphous alloys with higher specific strength of 300 to 400 MPa/(Mg/m^3) and higher specific modulus of 20 to 30 GPa/(Mg/m^3). The specific strength and specific modulus of the extruded bulk $Mg_{85}Ca_5Al_{10}$, $Mg_{97}Zn_1Y_2$, and $Mg_{96}Zn_1Y_3$ alloys were compared with the data of conventional Al-, Mg-, Ni-, and Ti-based alloys. The newly developed Mg-based alloys had higher specific strength of 270 to 330 MPa/(Mg/m^3) and high specific modulus of 19 to 23 GPa/(Mg/m^3). It is noticed that the specific strength values of the new Mg-based alloys are much superior to those for all the conventional metallic materials. RS P/M $Mg_{97}Zn_1Y_2$ alloy also exhibits high elevated temperature strength above 400 MPa in the temperature range up to 473 K and the further increase in testing temperature to 523 K causes a significant decrease in $\sigma_{0.2}$ to 130 MPa accompanying a significant increase in δ to 135%. It is also confirmed that the $\sigma_{0.2}$ values in the temperature range up to 473 K are much superior to those for the conventional heat resistant type Mg-based alloys of WE54-T6 and TMT-WE43.

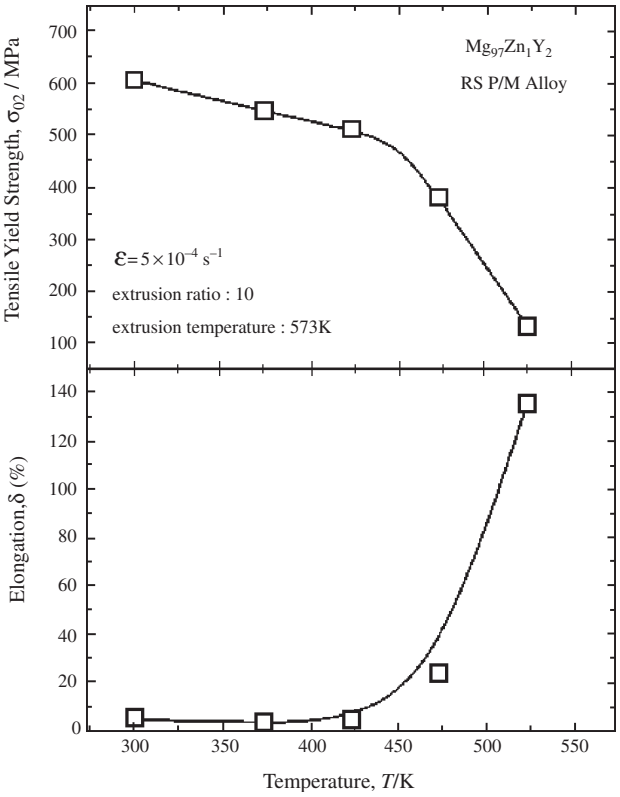


Figure 18. $\sigma_{0.2}$ and δ values as a function of extrusion temperature for the RS P/M $Mg_{97}Zn_1Y_2$ alloy. Reprinted with permission from [173], Y. Kawamura et al., *Mater. Trans.* 42, 1172 (2001). © 2001, Japan Institute of Metals.

The choice of Mg-rich compositions in the Mg–Zn–Y system causes a simultaneous increase in strength and ductility presumably because of the formation of a novel long periodic hexagonal matrix phase with much smaller grain sizes of 100 to 150 nm and the decrease in volume fraction of another phase. Figure 19 shows the TEM image of the RS P/M $Mg_{97}Zn_1Y_2$ alloy obtained by extrusion at 573 K. The alloy consists of fine grains with grain sizes of 100 to 150 nm. In addition, each grain contains a high density of stacking faults as well as very fine precipitates with a size of about 10 nm. In the selected-area electron diffraction pattern shown in Figure 19, one can see extra reflection spots at the positions of $1/3(0001)_{Mg}$ and $2/3(0001)_{Mg}$, indicating the formation of a long-periodic structure in which the periodicity is longer than that for the ordinary hcp-Mg phase.

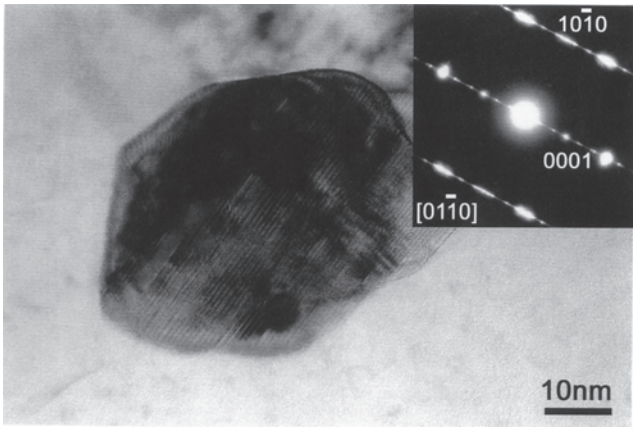


Figure 19. TEM image of the RS P/M $Mg_{97}Zn_1Y_2$ alloy obtained by extrusion at 573 K. Reprinted with permission from [173], Y. Kawamura et al., *Mater. Trans.* 42, 1172 (2001). © 2001, Japan Institute of Metals.

With the aim of clarifying in more details the long-periodic hexagonal structure, high-resolution TEM was used. One can see in Figure 20 an atomic configuration of ABACAB type which is three times longer than the ABABAB type atomic configuration [177] which is typical for the ordinary hcp-Mg phase. A novel long-period ordered structure is formed in the alloy, whose unit cell is composed of six close-packed planes of the magnesium crystal with a stacking sequence of ABCBCB′ where A and B′ layers are significantly enriched by Zn and Y. However, an atomic structure model of the Mg–(Zn, Y) long-period ordered structure was constructed based on a monoclinic lattice with $\gamma = 88°$, $a = 0.56$, $b = 0.32$, $c = 1.56$ nm, by discerning the lattice distortion from the ideal 6H-type structure. The lattice is distorted from an ideal hexagonal lattice of 6H-type (ABCBCB), which is probably due to an asymmetry of the chemical order with respect to the 6H-type stacking order.

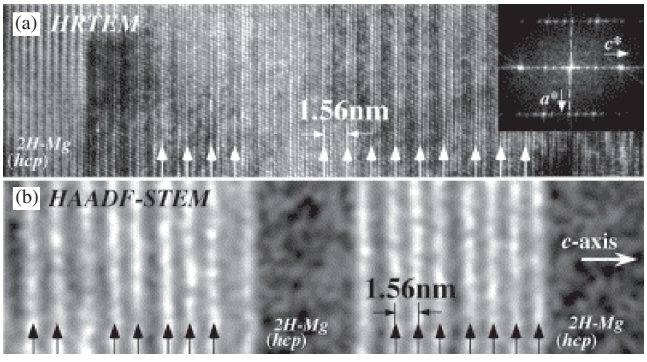


Figure 20. (a) HRTEM and (b) HAADF-STEM images showing the distribution of the long-period ordered structure in the fine-lamellar grain. HRTEM image of (a) reveals a periodic contrast with an interval of six close-packed planes (1.56 nm) at some places. Fourier transform obtained from the periodic contrast region is inserted in (a). HAADF-STEM image of (b) exhibits significant lamellar *Z*-contrasts with the same interval of 1.56 nm at some places, directly representing a chemical order of the structure with this correlation length. Reprinted with permission from [177], E. Abe et al., *Acta Mater.* 50, 3845 (2002). © 2002, Elsevier Science.

The present results demonstrate that the additional elements of a few atomic percent to Mg lead to formation of a long-period chemical-ordered as well as stacking-ordered structure. It is therefore concluded that the Mg–Zn–Y solid solution has a novel atomic configuration which is significantly different from that for hcp-Mg phase.

In addition to the change in the structure of the matrix phase, homogeneous dispersion of fine precipitates with cubic lattice and a size of about 7 nm was observed. These precipitates were characterized as a $Mg_{24}Y_5$ phase with a cubic structure of a lattice parameter of 1.1 nm, on the basis of the data on the high-resolution TEM image, the nanobeam electron diffraction pattern, and the nanobeam EDX profile [177].

In an equilibrium phase diagram in a Mg–Zn–Y system [178], there is no appreciable solid solubility limit of Zn and Y elements in hcp-Mg phase at room temperature. Consequently, the present Mg phase can be identified to be a reinforced solid solution saturated with Zn and Y elements. Such a reinforced solid solution generates significant strains because of significant atomic size ratios of 14% for Y/Mg and 20% for Mg/Zn. The high density of precipitates and stacking faults leads to significant strengthening. The mechanisms for the high strength and high ductility for the RS P/M Mg–Zn–Y alloys are thought to result from the combination of the following four factors: (1) nanograin hexagonal structure with grain sizes of 100 to 150 nm, (2) reinforced solid solution of Y and Zn elements with significantly different atomic sizes, (3) existence of a high density of stacking faults in the hexagonal Mg phase leading to significant change in the atomic configurations, and (4) homogeneous precipitation of cubic $Mg_{24}Y_5$ particles with a size of about 10 nm in the hexagonal Mg matrix.

5.1.3. Bulk Nanostructured Alloys Containing Quasicrystals

Quasicrystals not only attract interest themselves as a new type of solid. Being produced in the shape of nanoparticles with common boundaries or embedded in the amorphous matrix, they are used for strengthening of bulk samples and enhancing ductility of the metallic glasses. For example, a large number of such alloys have been recently produced in Al, Zr, and Hf based systems [44].

The icosahedral phase can contain dislocations [179]. However, it is very difficult for dislocations to move in the icosahedral phase because the movement of the dislocations demolishes the quasiperiodic lattice at room temperature and elevated temperatures. A single icosahedral $Al_{70}Pd_{10}Mn_{20}$ quasicrystal [180] prepared by the Chochralskii method has a high Young's modulus of 200 GPa, a high rigid modulus of 72 GPa, and a high Vickers hardness of about 750 at room temperature, even though it is extremely brittle [181]. The quasicrystal also exhibits high elevated temperature compressive strength of about 650 MPa at 1000 K, accompanying a significant elongation. By utilizing the high Young's modulus, high hardness, and high elevated temperature strength, an icosahedral base Al alloy with high strength were developed. The key point in the development is attributable to the possibility of synthesizing an icosahedral base alloy with good ductility.

High strength and good ductility for the icosahedral base alloys can be obtained in a nanostructure state where the icosahedral phase coexists with fcc α-Al phase and has a spherical morphology, small grain size, and homogeneous dispersion. The nanostructure state is also expected to be obtained by the increase in the stability of supercooled liquid against the transition to a crystalline phase. Significant attention has been paid to Al–Mn–RE and Al–Cr–RE systems because Al–Mn and Al–Cr alloys are well-known icosahedral-forming systems and Al–RE binary alloys are a good glass-forming system. The mixed structure consisting of icosahedral + α-Al phases is formed in the Al-rich melt-spun Al–Mn–RE alloys in the composition range above 92 at% Al and the mixed phase alloys have good bending ductility and high fracture strength reaching 1320 MPa. The high-strength $Al_{92}Mn_6Ce_2$ alloy consists of spherical icosahedral particles of about 50 nm size surrounded by a fcc-α-Al layer with a thickness of 5 to 10 nm. The halo-like reflection rings result from the icosahedral particles and the particles have random orientation because of the primary precipitation of the icosahedral phase, followed by the solidification of Al phase from the remaining liquid. The volume fraction of the icosahedral phase is as high as about 60 to 70%. A similar nanoscale mixed structure is also formed in rapidly solidified Al–Cr–Ce–Co alloys [182]. In the structure of $Al_{94.5}Cr_3Ce_1Co_{1.5}$ alloy the icosahedral particles with a size of 20–40 nm are homogeneously dispersed in the Al phase. The icosahedral phase alloy also exhibits a high fracture strength of 1340 MPa. The similar high fracture strength values exceeding 1000 MPa combined with high Vickers Hardness (H_v) above 400 are also obtained for rapidly solidified $Al_{93.5}Cr_3Ce_1Co_{1.5}TM_1$ alloys containing various transition metals [183]. The icosahedral phase based alloys also have good cold deformability and can be cold rolled up to a significant reduction ratio in thickness above 70% [184]. Besides, the cold-rolled alloy has a much finer mixed structure consisting of an icosahedral phase with a size of 5 to 10 nm coexistent with nanoscale fcc α-Al phase.

Here, it is important to point out that the high tensile strength combined with good ductility is obtained only for the mixed phase alloys prepared through the unique solidification process in which the nanoquasicrystalline phase precipitates as a primary phase, followed by the solidification of Al from the remaining liquid. On the other hand, when the Al-based alloys have the other type of solidification processes in which an Al solid solution is formed, followed by the precipitation of an icosahedral phase within Al grains and along Al grain boundaries or in which primary fcc-Al particles precipitate, followed by the solidification of an intergranular quasicrystalline phase from, the remaining liquid the alloys cannot have high tensile strength and good ductility.

In addition the same mixed structure of nanoscale icosahedral phase surrounded by fcc α-Al phase is also formed for atomized powders in Al–Mn–RE and Al–Cr–RE–TM systems as well as in Al–Mn–TM and Al–Cr–TM systems without RE elements [185]. When the atomized powders are extruded in the temperature range below 673 K, which is lower than the decomposition temperature of the icosahedral phase, one can obtain fully dense bulk icosahedral base alloys which have the same mixed structure as that for the atomized powders. Figure 21 shows the relation between tensile strength and plastic elongation for the extruded bulk icosahedral base alloys [154, 185, 186] in comparison with the data of conventional Al-based alloys obtained from the JIS handbook. The bulk icosahedral-based alloys have high σ_f values of 500 to 850 MPa combined with large elongation of 5–25% which are much superior to those for the conventional Al-based alloys. Furthermore, the bulk icosahedral base alloys can be divided into the following two kinds: (1) high-strength type for Al–Mn(Cr) alloys containing RE elements and (2) high ductility type for Al–Mn(Cr)–TM alloys without RE elements. The Al–Cr and Al–Mn base alloys containing RE elements also exhibit rather high elevated temperature strength of 350 MPa at 473 K and 200 MPa at 573 K, in addition to a high Young's modulus of about 100 GPa at room temperature. The bulk icosahedral based $Al_{94}Cr_{2.5}Co_{1.5}Mm_1Zr_1$ alloy also exhibits high fatigue strength values of 240 MPa at room temperature and 210 MPa at 423 K which are much higher than those (175 and 125 MPa) for conventional Al-based alloys in Al–Si–Fe systems. On the other hand, the bulk icosahedral base alloys without RE elements in $Al_{93-95}(Cr,Mn,Ni)_{5-7}$ and $Al_{93-95}(Cr,Mn,Cu)_{5-7}$ systems exhibit large elongation reaching 30% and high impact fracture energy reaching 160 kJ/m^2 combined with high tensile strength of 500 to 800 MPa [187]. All the values of tensile strength, elongation, and impact fracture energy are much higher than those for the 7075-T6 alloy.

More recently, a mostly single icosahedral phase was obtained at a new composition of $Al_{84.2}Fe_{7.0}Cr_{6.3}Ti_{2.5}$ in the melt-spun state [188]. There have been no data on the formation of an icosahedral phase in Al–Fe base alloys and hence the present formation of the mostly single icosahedral phase is believed to be the first evidence. Based on the previous data [189] that all the solute elements of Fe, Cr, and Ti have very low diffusivity in the fcc-Al phase, the

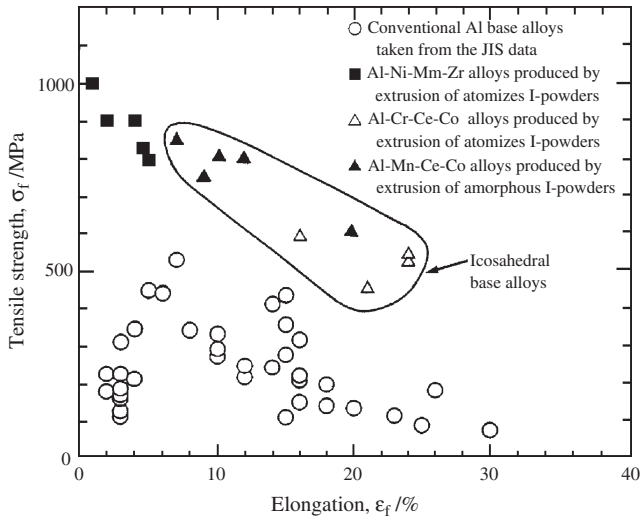


Figure 21. Relation between tensile strength and plastic elongation for the extruded bulk icosahedral base alloys. The data of conventional Al-based alloys are also shown for comparison. Reprinted with permission from [44], A. Inoue and H. M. Kimura, *J. Light Metals* 1, 31 (2001). © 2001, Elsevier Science.

icosahedral base alloys are expected to exhibit higher elevated temperature strength. Even for an $Al_{93}Fe_3Cr_2Ti_2$ alloy with a further Al-rich concentration, a similar mixed structure consisting mainly of icosahedral + Al is formed in the atomized powders with a particle size fraction range below 125 μm. The resulting bulk alloys prepared by extrusion of the atomized powders also consist mainly of icosahedral + Al phases. The new Al–Fe–Cr–Ti alloys exhibit excellent elevated temperature strength of 400 to 460 MPa at 473 K and 350 to 360 MPa at 573 K which exceed the air-force goal level. It has further been confirmed that the high elevated temperature strength level of the extruded $Al_{93}Cr_3Fe_2Ti_2$ alloy is maintained even after annealing for 1000 h at 573 K. The annealed alloy also keeps a fine mixed structure consisting of fine icosahedral particles with sizes less than 400 nm and fine Al grains with a size of about 500 nm.

Thus, the icosahedral based Al alloys can be classified into three groups: (1) high-strength type alloys in Al–Mn–RE and Al–Cr–RE–TM systems, (2) high ductility type alloys in Al–Mn–Cu–TM and Al–Cr–Cu–TM systems, and (3) high-elevated temperature strength type alloys in Al–Fe–Cr–Ti system. When the mechanical properties are compared with those for the 7075-T6 alloy (Fig. 22), one can notice that the present nanoquasicrystalline Al alloys exhibit a much better combination for almost all properties including tensile strength, elevated temperature strength, Charpy impact fracture energy, cold workability, elongation, and Young's modulus except thermal expansion. The excellent characteristics are quite attractive for future applications as a new type of Al-based alloy with high fracture strength, high ductility, and high elevated temperature strength.

Fine-grained magnesium alloys reinforced by quasicrystalline particles were developed by thermomechanical processes for as-cast $Mg_{96}Zn_{3.4}Y_{0.6}$ and $Mg_{94.8}Zn_{4.3}Y_{0.7}Zr_{0.2}$ alloys [190]. The $Mg_{96}Zn_{3.4}Y_{0.6}$ alloy shows a yield stress of 210 MPa and an ultimate tensile strength of 355 MPa, similar to those for an AZ91 alloy, and a large elongation of around 23%. The yield strength of Mg–Zn–Y alloys is relatively high due to the strengthening effect of the quasicrystalline particles [191]. However, at high temperature the level of flow stress of Mg–Zn–Y alloys is lower than that of commercial magnesium alloys due to the softness of the eutectic and fine grains. Furthermore, icosahedral phase particles in the Mg–Zn–Y alloy are stable against coarsening during deformation near the melting temperature of the eutectic, forming a stable quasicrystalline particle/matrix interface. This stability provides a large elongation with no voids opening at the interface between the quasicrystalline particle and the Mg matrix.

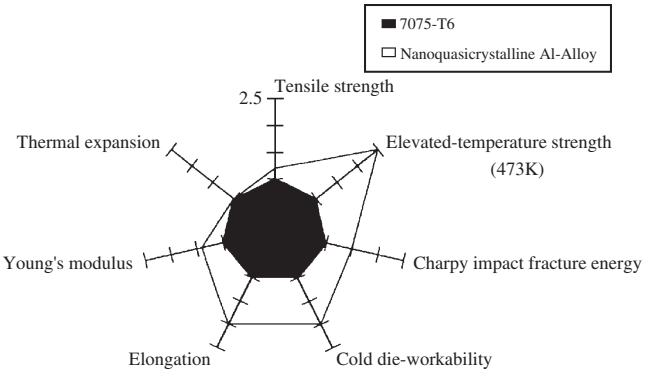


Figure 22. Comparison of mechanical properties of the icosahedral base bulk alloys with those for the commercial 7075-T6 alloy. Reprinted with permission from [44], A. Inoue and H. M. Kimura, *J. Light Metals* 1, 31 (2001). © 2001, Elsevier Science.

5.1.4. Nanostructured Alloys with Glassy Phase: Nanocomposites

As has been mentioned such materials can be produced using a single-step process, for example on rapid solidification by controlling cooling rate, or using a double-step one like devitrification of the glassy matrix.

Ductility of bulk glassy alloys can be drastically improved by precipitates of nanocrystalline or nanoquasicrystalline phase [119]. For example, in 1988 Al-based amorphous alloys with high tensile strength exceeding 1200 MPa were produced [192]. Subsequently, it has been noticed [193] that the homogeneous dispersion of nanoscale fcc α-Al particles in an amorphous matrix causes a drastic increase in tensile fracture strength to 1560 MPa which is three times higher than the strength level of conventional high-strength type Al-based alloys. For example as well as in the case of the ternary Al–Y–Ni alloys fine precipitates of fcc-Al formed by devitrification of the $Al_{85}Y_8Ni_5Co_2$ glassy matrix strengthen the alloy without reduction of bend ductility [194]. An optimum strength value was obtained for a crystalline phase volume fraction ranging from 5 to 25%. These nonequilibrium Al-based alloys exhibit much better mechanical properties compared with conventional crystalline alloys developed to date [195]. The volume fraction of the fcc-Al phase is controlled by changing the rotation speed of wheel and is evaluated by the change in the exothermic heat release due to crystallization. The tensile fracture strength increases from about 1100 MPa at volume fraction (V_f) of fcc-Al V_f = 0% to 1560 MPa at $V_f = 25\%$, accompanying an increase in Vickers hardness from 280 to 400 and Young's modulus from 63 to 71 GPa [196]. The significant decrease in tensile fracture strength by the further increase in V_f is due to the embrittlement of the remaining amorphous phase by the progress of structural relaxation and the enrichment in the solute elements [197]. A similar nanoscale mixed structure is also formed by annealing in the first-stage exothermic reaction range and the mixed phase alloys exhibit high tensile strength exceeding 1400 MPa which is about 1.4 times higher than that for the corresponding amorphous single phase alloys [198].

$Al_{88}Y_2Ni_9Fe_1$ amorphous alloy containing 7% volume fraction of the fcc-Al particles with a size of about 3–4 nm has a high tensile fracture strength of 1320 and 1260 MPa with volume fraction of 24%, respectively [199]. It is notable that these values for the former alloy are higher than the highest values of Al-based amorphous single-phase alloys. The high strength is combined with good bend ductility.

The deformation of amorphous materials is concentrated on the shear band of maximum shear stress which maintains about 45 degrees with the load (tensile or compressive) direction. The width of the shear band in Al alloys is about 20 nm.

Table 2. Alloy system, microstructure, and mechanical properties for the quasicrystalline (Q) phase bearing bulk Al-based alloy.

Type	Alloy system	Structure	Mechanical properties
High-strength alloy	Al–Cr–Ce–TM, Al–Mn–Ce	Al + Q	$\sigma_f \sim$ 600–800 MPa, $\varepsilon_p \sim$ 5–10%
High ductility alloy	Al–Mn–Cu–TM, Al–Cr–Cu–TM	Al + Q	$\sigma_f \sim$ 500–600 MPa, $\varepsilon_p \sim$ 12–30%
High elevated temperature strength alloy	Al–Fe–Cr–Ti	Al + Q + $Al_{23}Ti_9$	$\sigma_f \sim$ 350 MPa at 573 K

Source: Reprinted with permission from [44], A. Inoue and H. M. Kimura, *J. Light Metals* 1, 31 (2001). © 2001, Elsevier Science.

An $Al_{88}Ni_{10}Ce_2$ glassy alloy primarily transformed from an amorphous to an amorphous plus fcc α-Al state exhibited a large elongation reaching 45% in the temperature range of 450 to 465 K. In this temperature range primary fcc-Al particles with a size of about 12 nm precipitate homogeneously in the amorphous matrix with an interparticular spacing of about 6 nm. The appearance of the significant elongation is due to a crystallization-induced elongation phenomenon. The process occurs through inhomogeneous deformation along shear plane, followed by deformation-induced preferential precipitation of Al particles on the shear plane and then propagation of the deformation site to other site resulting from the precipitation-induced strengthening in the deformed area [200].

The increase in strength by the dispersion of nanoscale fcc-Al particles in the amorphous matrix has been thought [44] to result from the combination of the following three effects: (1) the defect-free effect of nanoscale fcc-Al particles, because the Al particles are too fine to contain dislocations, (2) the interface effect, because the amorphous/Al particle interface has a highly dense packed structure without excess vacancies due to a low interfacial energy, and (3) the nanoscale effect, because the Al particle size is smaller than the width of the inhomogeneous shear deformation region and hence the nanoscale Al particles can act as an effective barrier against the shear deformation of amorphous matrix [201]. It also explains maintenance of a good bend ductility.

At the same time, it was also suggested [202] that the hardening could be attributed mainly to solute enrichment of the residual glassy matrix due to lowering of the Al content. This theory well describes hardening of the material. As the hardness of fully amorphous or partially crystalline alloys correlates well with solute content in the amorphous phase it is suggested [203] that not only Al particles but also the amorphous matrix has some role in the hardening and embrittlement of the alloy.

A model of the ductile–brittle transition of partially crystallized amorphous Al–Ni–Y alloys was proposed in [204]. The maximum strength of the partially crystalline Al–Ni–Y alloys was obtained when the size of Al particles is 10 nm. However, partially crystallized Al–Ni–Y alloys become brittle when the size and the volume fraction of the Al particles exceeded the optimum values. The reactions in three regions, Al particles, Y accumulated interface layers, and amorphous matrix, were derived by balancing the quantity of each element. It is assumed that the Al–Ni–Y alloys become brittle when the solute content of the interface layers reached the critical value (20%). The predictions were found to be in good agreement with the published data. However, the tensile strength is too high to get a consolidated bulk alloy by warm extrusion in a relatively low temperature range where the nanoscale mixed structure is maintained.

Bulk $Al_{88.5}Ni_8Mm_{3.5}$ alloy having a mixed structure consisting of amorphous and fcc-Al phases was produced by extrusion of atomized powders at an extrusion ratio of 10 and at 633 K and was found to exhibit high tensile strength of 940 MPa at room temperature, 700 MPa at 423 K, and 520 MPa at 473 K [205]. Similarly, the yield strength is also as high as 845 MPa at room temperature and 420 MPa at 473 K. The high values of strength remain unchanged even after annealing for 100 h at these testing temperatures. The fatigue limit after the cycles of 107 is also as high as 330 MPa at 293 K and 220 MPa at 473 K. The fatigue limit at room temperature is also 1.2 times as high as the highest value for conventional Al-based alloys.

In the historical progress of the high-strength Al- and Zr-based alloys, a drastic increase in tensile strength is recognized for the recent nonequilibrium alloys including an amorphous phase (Figs. 14 and 15).

However, in some cases formation of the primary α-Al particles worsens the mechanical properties of Al-based alloy. For example, partial replacement of Ni by Cu in the $Al_{85}Y_8Ni_5Co_2$ metallic glass caused formation of the nanoscale α-Al particles that resulted in sharp drops in tensile strength, hardness, crystallization temperature, and disappearance of the supercooled liquid prior to crystallization (Fig. 23) [206]. Cu has a much lower absolute value of heat of mixing with Al, Y, and Co than Ni [207] which leads to a decrement of the interatomic constraint force. Thus, Cu may weaken the interaction needed for the stability of the glass, thus resulting in the disappearance of T_g and precipitation of α-Al nanocrystals. In addition the volume fraction of the α-Al nanocrystals in $Al_{85}Y_8Ni_3Co_2Cu_2$, for example, is much lower than that in the primarily devitrified $Al_{85}Y_8Ni_5Co_2$ metallic glass [208, 209]. α-Al interparticular distances in the $Al_{85}Y_8Ni_3Co_2Cu_2$ metallic glass significantly exceed the particle size (Fig. 24). Thus, α-Al particles cannot act as an effective barriers against the shear deformation of amorphous matrix as was supposed for the $Al_{85}Y_8Ni_5Co_2$ metallic glass.

At the same time existence of the MRO zones and nanoscale particles in as-cast rod-shaped bulk (2 mm in diameter) $Zr_{55}Co_{20}Al_{12}Cu_5Nb_8$ glassy alloy (see Fig. 16) enhances its ductility leading to appearance of a plastic deformation region of about 1% deformation as shown in Figure 25.

It can be shown that the cylindrical bulk sample of $Zr_{65}Al_{7.5}Cu_{7.5}Ni_{10}Pd_{10}$ alloy, the structure of which consists of an almost single icosahedral phase with nanoscale grain sizes of 20–40 nm, shows a better combination of

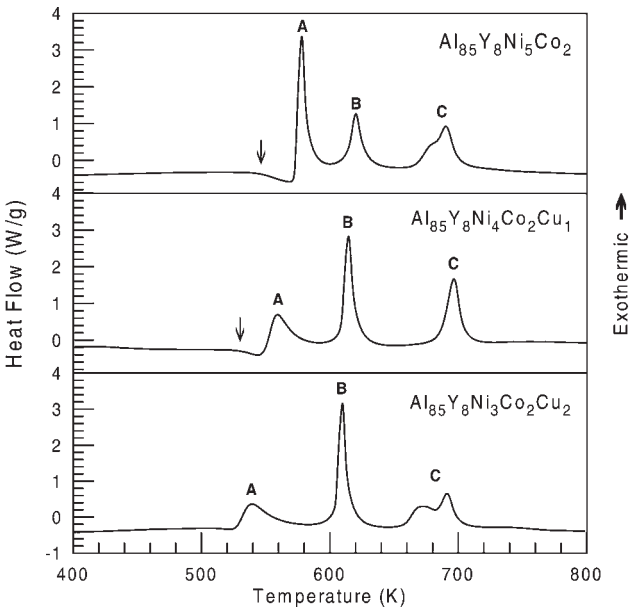


Figure 23. DSC traces of $Al_{85}Y_8Ni_{5-3}Co_2Cu_{0-2}$ metallic glasses taken at 0.67 K/s. The arrows indicates increment of C_p related to the glass transition. Reprinted with permission from [206], D. V. Louzguine and A. Inoue, *J. Mater. Res.* 17, 1014 (2002). © 2002, Materials Research Society.

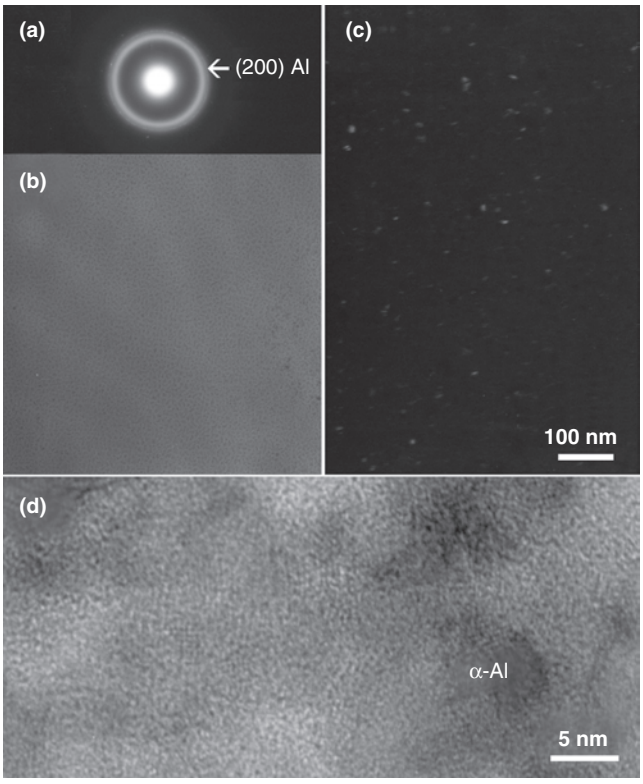


Figure 24. TEM of $Al_{85}Y_8Ni_3Co_2Cu_2$ metallic glass in as-solidified state. (a) Selected-area electron diffraction pattern, (b) bright-field (c) dark-field, and (d) high-resolution image. Reprinted with permission from [206], D. V. Louzguine and A. Inoue, *J. Mater. Res.* 17, 1014 (2002). © 2002, Materials Research Society.

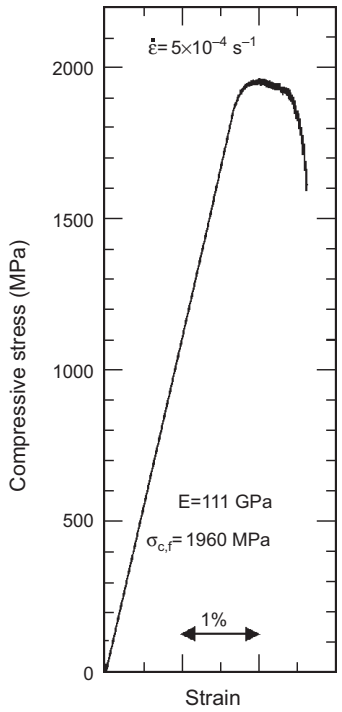


Figure 25. Compressive stress–strain curve of the rod-shaped bulk (2 mm in diameter) $Zr_{55}Co_{20}Al_{12}Cu_5Nb_8$ glassy alloy containing MRO zones and nanoscale particles. Reprinted with permission from [120], T. Wada, Master's Thesis, Tohoku University, 2003.

mechanical properties compared to the as-cast glassy sample. The Young's modulus, ultimate tensile strength, 0.2% proof stress, and total percentage elongation including elastic elongation are 85 GPa, 1750 MPa, 1640 MPa, and 2.2%, respectively, for the glassy alloy and 88 GPa, 1780 MPa, 1830 MPa, and 3.1%, respectively, for the icosahedral phase-based alloy [210]. However, it is well known that stoichiometric icosahedral phase-based alloys have high hardness and are extremely brittle [211] and it is difficult to consider that the icosahedral phase itself in the present alloys has plastic deformability. The good mechanical properties are attributed to the residual existence of an intergranular glassy phase, though the volume fraction of the glassy phase is estimated to be less than 5%. The deformation of the icosahedral-based alloy takes place in the intergranular glassy phase along the maximum shear plane. The existence of the icosahedral particles can act as a resistant medium against the shear deformation (Fig. 26). The good ductility of the glassy layer is also attributed to the localization effect of deformation. The localization generates the multiaxis stress condition as well as the increase in temperature.

A nanocrystalline bulk $Zr_{55}Al_{10}Cu_{30}Ni_5$ alloy containing nanoscale crystals embedded uniformly in a glassy matrix also has both a high tensile strength of 1.7 GPa and good ductility [212].

Deformation behavior of partially crystallized Zr-based glassy alloys consisting of a metallic glass matrix with crystallite precipitates of different shape and size was studied [213].

However, although a finely dispersed structure consisting of a glassy phase and nano-submicron-sized $Ni_{33}Cr_{33}P_{34}$

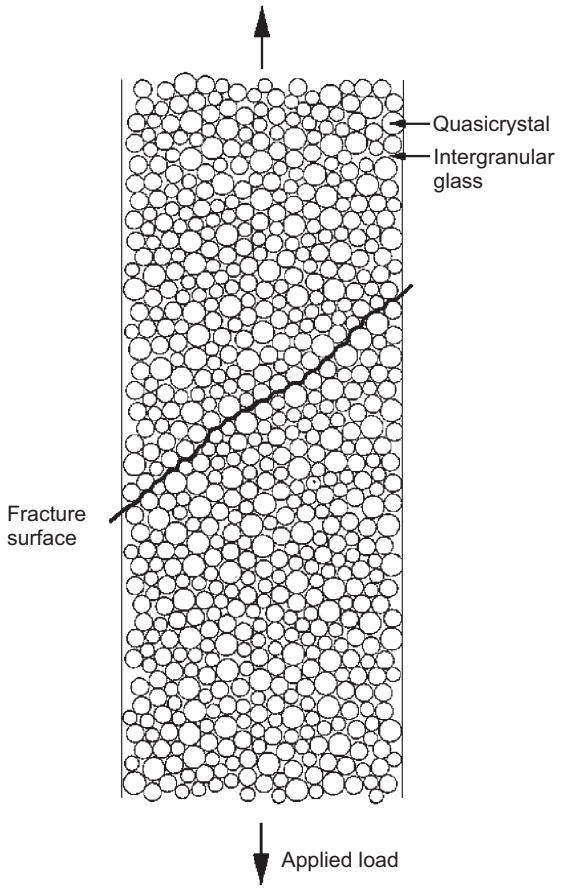


Figure 26. Schematic illustration of the deformation and fracture modes of the nanoquasicrystalline Zr–Al–Ni–Cu–Pd alloy. Reprinted with permission from [25], A. Inoue et al., *Sci. Technol. Adv. Mater.* 1, 73 (2000). © 2000, Elsevier Science.

crystalline particles (≈500 nm in size) with a volume fraction of 9–13% was formed in $Pd_{40}Cu_{29}Ni_{10}P_{20}Cr_1$ alloy prepared by copper-mold casting [214] its mechanical strength is lower than that of $Pd_{40}Cu_{30-x}Ni_{10}P_{20}$ glassy alloy. This result supports the proposed strengthening mechanism of nanocomposite materials by blocking of deformation share bands of 20–50 nm in width depending upon alloy composition in glassy matrix by the precipitated nanoparticles. Particles of 500 nm are too large to be the barriers blocking deformation.

Bulk composite Zr-based alloys consisting of a $Zr_{55}Cu_{30}Al_{10}Ni_5$ glassy matrix reinforced with micrometer size ZrC particles exhibited high strength and good ductility [215, 216]. The Young's modulus, compressive strength, and Vickers hardness of the composite alloys increase almost linearly from 103 to 125 GPa and 1820 to 2170 MPa and 488 to 563, respectively, with increasing volume fraction of ZrC from 0 to 17 vol%. The plastic elongation also increases from nearly zero at 0 vol% of ZrC to 0.5% at 10 vol% of ZrC.

5.1.5. Superplasticity

Superplastic flow in an alloy is achieved at a sufficiently small-micrometer and submicrometer grain size [217]. Recently this phenomenon has been extended to nanomaterials. The large volume fraction of grain boundaries and their small size are believed to enhance superplastic behavior.

The Al–Ni–Mm alloy with nanocrystalline structure has been found [218] to lead to a high-strain-rate superplasticity which is evidenced from a high-strain-rate sensitivity exponent (m value) of 0.5 as well as a large elongation reaching about 670% at a high strain rate of 1.0 s^{-1}.

A nanocrystalline Ni_3Al with a grain size of about 50 nm processed by severe plastic deformation exhibited superplasticity under tensile test at 923–998 K and a strain rate of 1×10^{-3} s^{-1}. The total elongations for as-processed specimens were 380 and 560% at 923 and 998 K, respectively, which is really impressive for the intermetallic phase. Nanocrystalline Ni_3Al showed quite low reduced superplastic temperature, a large strain hardening stage, and high flow stresses. The tensile elongation of 560% at 998 K was obtained, whereas 640% elongation at nearly the same strain rate was reported for the polycrystalline Ni_3Al with a 6 μm grain size doped with boron [219] when the testing temperature was 1373 K, that is, more than 300 K higher.

The mean grain size of nanocrystalline Ni_3Al after deformation increased up to about 100 nm while no trace of dislocation activity was found by TEM. It is suggested that the deformation took place by intergranular sliding. The role of transformations of grain boundary dislocation pile-ups at triple junctions of grain boundaries in plastic deformation processes in nanocrystalline materials with special attention paid to competition between different deformation mechanisms has been discussed in [220].

High elongation above 175% was obtained in nanosubmicron-grained Ti-45Al-2.4 Si alloy at 800 K [221, 222]. High-strain-rate superplasticity was also obtained in nanocrystalline Al-based 1420 alloy [223].

Low-temperature superplasticity has been reported in nanocrystalline Ni, nanocrystalline Al-based alloy (1420-Al), and nanocrystalline Ni_3Al [224]. Nanocrystalline electrodeposited Ni exhibited an elongation of 295% achieved at 623 K which is 0.36 of the melting temperature of Ni. A maximum elongation of 895% at a strain rate of 10^{-3} s^{-1} was obtained at a temperature of 693 K [225].

RS P/M $Mg_{97}Zn_1Y_2$ alloy also exhibited the high strain rate superplasticity as is evidenced from large δ above 700% at 623 K in the vicinity of the high strain rate of 0.1 s^{-1} as well as high strain rate sensitivity exponent of 0.4. It is recognized that the sample with the largest elongation at the high strain rate of 0.10 s^{-1} was elongated through a homogeneous deformation mode. Figure 27 shows the relation between elongation (δ) and strain rate for the RS P/M $Mg_{97}Zn_1Y_2$ alloy. The δ value is the highest for the Mg–Zn–Y alloy and the strain rate leading to the highest δ value is also located at the high strain rate side. Thus, the RS P/M $Mg_{97}Zn_1Y_2$ alloy has excellent mechanical properties as well as high strain rate superplasticity [226].

The Mg–Ca–Al alloy also exhibits a very large elongation above 700% at 723 K and at a very high strain rate of 0.2 s^{-1} [39, 227]. It is also reported [227] that a nearly linear relation exists between flow stress and strain rate. From the linear relation, the strain rate sensitivity exponent (m value) was evaluated to be 0.45 in the vicinity of 0.2 s^{-1}. These data

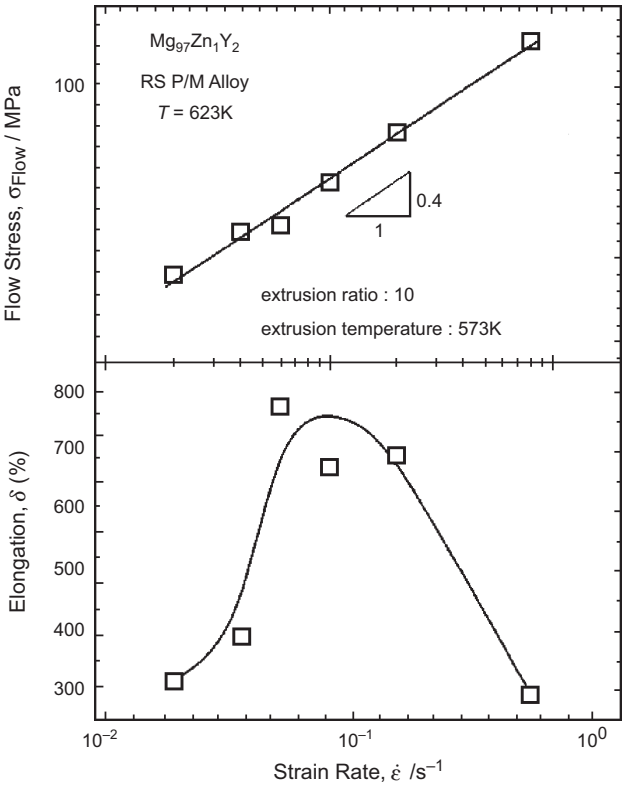


Figure 27. The relation between δ and strain rate for the RS P/M $Mg_{97}Zn_1Y_2$ alloy. Reprinted with permission from [173], Y. Kawamura et al., *Mater. Trans.* 42, 1172 (2001). © 2001, Japan Institute of Metals.

indicate the achievement of high strain rate superplasticity for the extruded Mg–Ca–Al bulk alloy.

The deformation behavior of a bulk $Cu_{47}Ti_{33}Zr_{11}Ni_6Sn_2Si_1$ bulk glassy alloy (Fig. 28) in the supercooled liquid region has been studied in [228]. It is suggested that the alloy devitrifies during deformation while no structure data except for the XRD are given in the as-solidified state. However, the shape of the X-ray diffraction pattern in Figure 28 assumes the possible existence of nanocrystals in an as-solidified sample. Two broad peaks in the XRD pattern in the as-cast state (Fig. 28) might be likely caused by nanocrystals. Nevertheless, the sheet shows large elongation around 450% at temperatures of 745 and 750 K (2×10^{-3} s^{-1}) and more than 650% at a strain rate of 1×10^{-2} s^{-1} (750 K) [228]. TEM images (Fig. 29) show nanocrystalline particle bands aligned along the loading direction.

5.2. Magnetic Properties

Magnetic material is another very important and dynamically developing field of the application of nanostructured metallic materials and the connection between nanostructure and magnetic properties is a topic of intensive investigations.

5.2.1. Soft Magnetic Alloys

Classical examples of Fe-based soft magnetic materials with mixed nanocrystalline and amorphous structure are Finemet $Fe_{73.5}Cu_1Nb_3Si_{13.5}B_9$ [229] and Nanoperm

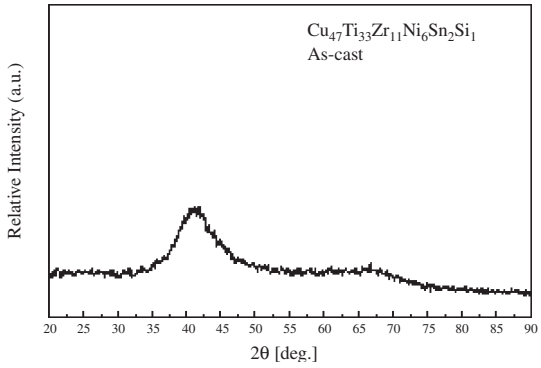


Figure 28. X-ray diffraction pattern of an as-cast $Cu_{47}Ti_{33}Zr_{11}Ni_6Sn_2Si_1$ alloy sheet. Reprinted with permission from [228], D. H. Bae et al., *Acta Mater.* 50, 1749 (2002). © 2002, Elsevier Science.

$Fe_{84}Zr_{3.5}Nb_{3.5}B_8Cu_1$ [230] alloys, the structure of which consists of bcc-Fe nanocrystals of below 20 nm in size finely dispersed in the amorphous matrix. Hitperm $Fe_{44}Co_{44}Zr_7B_4Cu_1$ [231] alloy has an α'-FeCo phase. Cu, Nb, or Zr, despite on their low content, are responsible for bcc-Fe grain

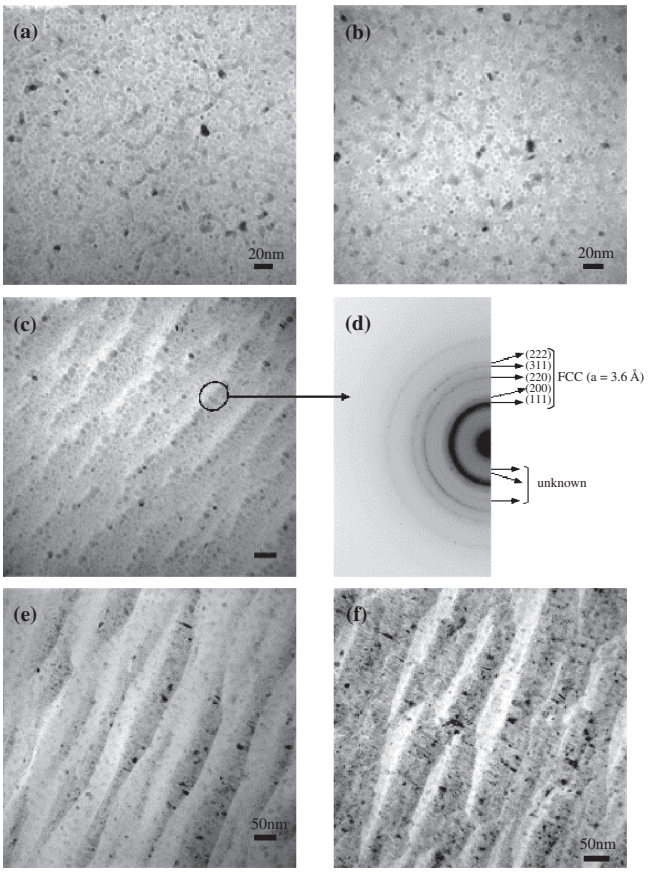


Figure 29. Bright-field TEM images of the alloy tested at a strain rate of 2×10^{-3} s^{-1} and a temperature of 750 K in which (a), (b), and (c) are taken from the specimen deformed to $\varepsilon = 0.5$ and (e) and (f) are taken from the specimen deformed to $\varepsilon = 1.0$. Also the electron diffraction pattern obtained from the crystals marked in (c) is given in (d). Reprinted with permission from [228], D. H. Bae et al., *Acta Mater.* 50, 1749 (2002). © 2002, Elsevier Science.

refinement and formation of a nanostructure in these alloys. The best results were obtained for alloying with two minor alloying elements like Cu and Nb or Cu and Zr (Fig. 30), for example.

Cu has a very low solid solubility in iron and as has been shown by means of atom probe field ion microscopy and high resolution transmission electron microscopy [233] forms nanoclusters in the $Fe_{73.5}Si_{13.5}B_9Nb_3Cu_1$ amorphous matrix which works as a heterogeneous nucleation site for bcc-Fe particles on devitrification. X-ray absorption fines structure studies also showed that Cu clusters with near-fcc structure were present from very early stages of the devitrification process [234]. The density of the clusters estimated by a three-dimensional atom probe is on the order of 10^{-24} m^{-3} at the average cluster size of about 2 nm [235]. Structure evolution in $Fe_{73.5}Si_{13.5}B_9Nb_3Cu_1$ alloy under primary devitrification is shown in Figure 31. Extended X-ray absorption fine structure analysis of grain boundaries in nanocrystalline $Fe_{85}Zr_7B_6Cu_2$ alloys showed very low Fe content in the grain boundaries between bcc-Fe nanograins.

These alloys show high permeability (μe) and low magnetic losses [236]. $Fe_{84}Zr_{3.5}Nb_{3.5}B_8Cu_1$ alloy shows a very high μe of 100,000 combined with the high magnetization (B_s) of 1.53 T [237]. This excellent μe is comparable to those of nanocrystalline $Fe_{73.5}Si_{13.5}B_9Nb_3Cu_1$ and the zero-magnetostrictive Co based amorphous alloys, and the high B_s is comparable to other Fe-based amorphous and nanocrystalline alloys with good soft magnetic properties.

It has also been found that the nanoparticles in partially crystallized $Fe_{73.5}Si_{13.5}B_9Nb_3Cu_1$ (Finemet) and $Fe_{92-x}B_xZr_{7-y}Nb_yCu_1$, with $x = 2$, 4, 6, 8, 10, 12 and $y = 0$, 3.5 at% (Nanoperm-type) nanostructured soft magnetic alloys show a superferro–superpara- and a superpara–paramagnetic transition at high temperature [238]. The transition temperatures depend on the annealing conditions.

Fe–Zr–B soft magnetic materials also exhibit high magnetization of 1.60–1.70 T under an applied field of 800 kA/m as well as high effective permeability (μe) of 13,000–15,000 at 1 kHz [183]. Typical nanocrystalline bcc $Fe_{89}Hf_7B_4$ and $Fe_{84}Nb_7B_9$ alloys subjected to the optimum annealing exhibit high magnetization above 1.5 T as well as high effective permeability (μe) at 1 kHz above 20,000. The bcc Fe–M–B (M = Zr, Hf, or Nb) alloys also show low core losses (W)

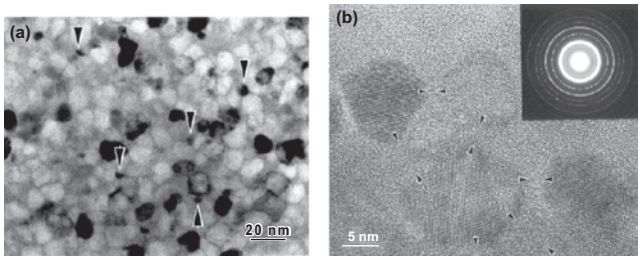


Figure 30. (a) Bright-field TEM image of the melt-spun $Fe_{73.5}Si_{13.5}B_9Nb_3Cu_1$ alloy annealed at 823 K for 60 min (Finemet). The grains indicated by arrows are fcc-Cu. (b) HRTEM image and inset ring pattern. The arrows indicate the residual amorphous phase surrounding the Fe nanocrystals. Reprinted with permission from [232], K. Hono, *Progr. Mater. Sci.* 47, 621 (2002). © 2002, Elsevier Science.

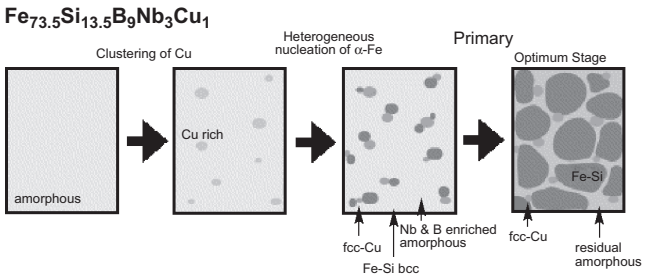


Figure 31. Schematic illustration of the microstructural evolution in melt-spun $Fe_{73.5}Si_{13.5}B_9Nb_3Cu_1$ amorphous alloy by primary crystallization. (Si is partially dissolved in Fe.) Reprinted with permission from [232], K. Hono, *Progr. Mater. Sci.* 47, 621 (2002). © 2002, Elsevier Science.

from 1.4×10^{-1} to 2.1×10^{-1} W/kg at 50 Hz and 1.4 T and from 1.70 to 2.50 W/kg at 1 kHz and 1.0 T [239].

Annealed $Fe_{82}(Zr,Hf,Nb)_7B_{10}Cu_1$ alloys exhibit good soft magnetic properties especially in the high-frequency range [240]. The high concentration of boron suppresses growth of bcc-Fe grains and stabilizes the amorphous matrix. The result reveals that the reason for good soft magnetic properties is associated with the reduction of the magnetic anisotropy by rather strong exchange energy when annealing temperature is increased.

Nanocrystalline $Fe_{42.5}Co_{42.5}Nb_7B_8$ alloy with a structure consisting of nearly spherical bcc grains with size from 5 to 10 nm dispersed in the residual amorphous matrix was obtained by annealing of the amorphous alloy at 873 K for 3.6 ks and exhibits a high saturation magnetization (I_s) of 1.90 T and a low coercivity (H_c) of 60 A/m [241]. It also exhibits a very high Curie temperature (T_c) more than 1173 K. The segregation of the Nb element in the intergranular amorphous phase increases the thermal stability of the amorphous phase and suppresses the grain growth of the bcc phase. The high thermal stability of the structure and the high Curie temperature are promising for soft magnetic materials for high temperature applications.

The origin of the good soft magnetic properties is as follows: (a) The apparent anisotropy is decreased by the combined effects of the formation of the nanoscale bcc structure and the achievement of rather strong magnetic coupling between the bcc grains through the intergranular ferromagnetic amorphous phase. (b) The small saturation magnetostriction (λ_s) results from the nonequilibrium bcc phase. The solute-rich intergranular amorphous phase with high Curie temperature (T_c) and high thermal stability plays an important role in the achievement of the good soft magnetic properties through the attainment of the rather strong magnetic coupling between the bcc grains.

5.2.2. Hard Magnetic Alloys

Two types of permanent magnetic materials consisting mainly of Fe_3B with $Nd_2Fe_{14}B$ phase were obtained by annealing melt-spun $Nd_{4.5}Fe_{77}B_{18.5}$ flakes [242]. The microstructure is composed of nanoscale grains of magnetically soft Fe_3B and magnetically hard $Nd_2Fe_{14}B$ phases. High remanence of 0.8 Ms is obtained due to the remanence enhancement effect of exchange-coupled magnetic

grains. $J_r = 1.2$ T and $(BH)_{max} = 97$ kJ/m^3 while the coercivity, $H_c = 240$ kJ/m, is slightly low. Influence of heating rate on the microstructure of $Fe_3B/Nd_2Fe_{14}B$ nanocomposite magnets has also been recently studied [243]. Attempts to improve the coercivity of this type of magnet were made [244]. With increasing Cr content in $Nd_{4.5}Fe_{77-x}Cr_xB_{18.5}$ alloy, the intrinsic coercivity (H_{cI}) increased; however, the remanence (B_r) decreased considerably [245]. High coercivity values exceeding 300 kA/m were obtained in amorphous $Nd_5Fe_{72}Cr_5B_{18}$ crystallized into $Fe_3B/Nd_2Fe_{14}B$ state [246].

Low coercivity of $Fe_3B/Nd_2Fe_{14}B$ magnets impedes a limit on their application. The bcc-$Fe/Nd_2Fe_{14}B$ nanomaterials have higher coercivity than those for $Fe_3B/Nd_2Fe_{14}B$ magnets. H_c of 480 kA/m and $(BH)_{max}$ of 160 kJ/m^3 are reported for bcc-$Fe/Nd_2Fe_{14}B$ nanomaterials. The structure consist of two phases: magnetically hard $Nd_2Fe_{14}B$ with nanoparticles of α-iron on grain boundaries. Grain size of the $Nd_2Fe_{14}B$ is below 30 nm and particle size of the α-iron is below 10 nm [247].

α-$Fe/Nd_2Fe_{14}B$ nanostructured magnet of $Fe_{89}Nd_7B_4$ composition contains residual amorphous phase and shows $B_r = 1.22$ T, $H_c = 240$ kA/m, and $(BH)_{max} = 130$ kJ/m^3. This alloy has high iron and low boron concentrations [248].

Rapidly quenched Sm–Fe–N powders show high remanence values. The structure consists of a nanoscale mixture of $SmFe_7N_x$ phase and α-Fe with a size of 20–30 nm [249]. High remanence is due to the small crystal size and exchange coupling at interphase boundaries.

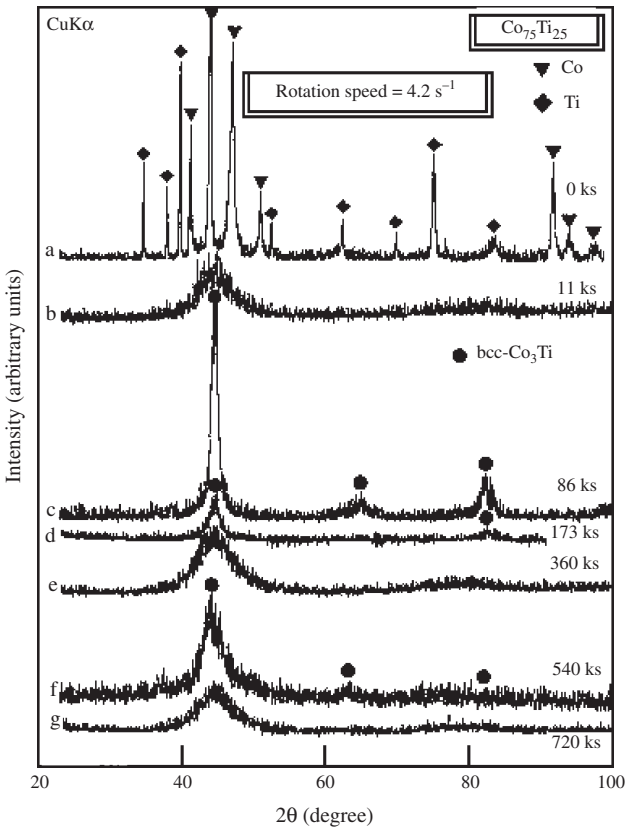


Figure 32. XRD patterns of ball-milled $Co_{75}Ti_{25}$ powders after (a) 0, (b) 11, (c) 86, (d) 173, (e) 360, (f) 540, and (g) 720 ks of milling time, using a rotation speed of 4.2 s^{-1}. Reprinted with permission from [250], M. Sherif El-Eskandarany et al., *Acta Mater.* 50, 1113 (2002). © 2002, Elsevier Science.

6. SOLID STATE PHASE TRANSFORMATIONS IN NANOSTRUCTURED ALLOYS

Most nanostructured materials, being metastable, transform on heating. In this section we will also show the examples of phase transformations in nanomaterials that have not been included in other sections.

6.1. Cycling Transformations

Cycling amorphous-to-crystalline phase transformations were observed in different binary TM–TM alloys [250]. For example, cycling phase transformations were observed during ball milling of elemental $Co_{75}Ti_{25}$ powder. After a few kiloseconds of milling, a solid-state reaction takes place at the fresh interfaces of Co/Ti layers and an amorphous phase of $Co_{75}Ti_{25}$ is formed after 11 ks (Figs. 32 and 33). This amorphous phase transforms into an ordered fcc-Co_3Ti upon heating to 880 K (crystallization). Further milling (86 ks) also leads to an amorphous–crystalline transformation (crystallization) and the formed phase is metastable bcc-Co_3Ti nanocrystalline powders. This transformation is attributed to the inability of the formed amorphous phase to withstand the impact and shear forces that are generated by the milling media (balls). When the milling time is increased (360 ks), the crystalline phase is subjected to several defects (points and lattice defects) that raise the free energy of bcc-Co_3Ti to a less stable phase (amorphous). Here, the crystalline–amorphous transformation which takes place is similar to the mechanical grinding method [251] in which the amorphization occurs by relaxing the short-range order without any compositional changes. Further milling leads to the formation of crystalline and/or amorphous phases depending on the milling time.

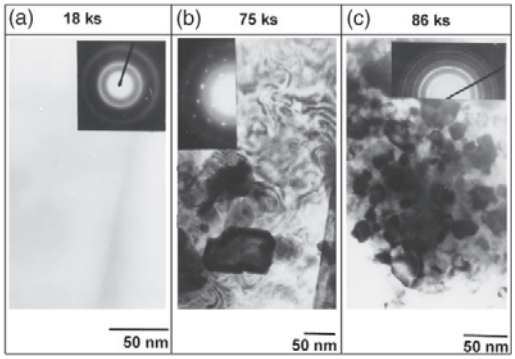


Figure 33. Bright-field TEM images and the corresponding selected-area electron diffraction patterns of as-ball-milled $Co_{75}Ti_{25}$ powders after (a) 18, (b) 75, and (c) 86 ks of milling time using a rotation speed of 4.2 s^{-1}. Reprinted with permission from [250], M. Sherif El-Eskandarany et al., *Acta Mater.* 50, 1113 (2002). © 2002, Elsevier Science.

6.2. Phase Transformation in Ge–Al–La Alloy

The structure of the as-solidified $Ge_{60}Al_{25}La_{15}$ alloy is a mixture of amorphous and metastable nanocrystalline phase [252]. It consists of colonies of needle-shaped crystalline phase with a nanoscale size in two dimensions crystallized within the amorphous matrix, as shown in Figure 34. Some nanoscale crystalline particles are also present in amorphous matrix, but their volume fraction is low.

The $Ge_{60}Al_{25}La_{15}$ alloy has a heating-rate-dependent transformation behavior [252]. Being heated using DSC at a heating rate of 0.33 K/s and higher this alloy showed the only single exothermic peak up to melting, whereas at lower heating rates of 0.17 and 0.08 K/s for example, as shown in Figure 35, about three heat effects are seen. It is important to note that total heat released during a single thermal effect of 85.7 J/g observed at a heating rate of 0.33 K/s is almost equal to total heat effects of 83.7 and 86.2 J/g observed at 0.17 and 0.08 K/s, respectively, calculated as a sum of all three effects [252].

The reaction related to the first DSC peak scanned at heating rates of 0.17 K/s or less was found to be "Amorphous phase + unknown crystalline phase 1 → unknown phase 2." This reaction is similar to the peritectic reaction in binary systems with amorphous phase instead of the liquid one. During subsequent two reactions pure *Fd3m* Ge precipitates from amorphous matrix and an unknown phase 2 transforms into unknown phase 3. This phase is stable in the $Ge_{60}Al_{25}La_{15}$ alloy. In accordance with equilibrium Ge–Al–La phase diagram it can be identified as Ge_2AlLa phase of unknown structure observed in [253].

The structure of the $Ge_{60}Al_{25}La_{15}$ alloy after the first strong heat effect changes greatly. The inhomogeneous as-solidified structure consisting of separate amorphous and crystalline phases (see Figs. 34 and 36a) transforms to a relatively uniform one (Figs. 36b and 37) in which needle-shaped particles of GeAlLa phase of about 10 nm in thickness are homogeneously distributed in the residual amorphous matrix. Due to the small crystal thickness diffraction peaks belonging to unknown phase 2 (marked by ▼ in Fig. 38) are broadened. Residual crystalline phase 1 is also seen in Figure 38.

Such a structural change takes place at a relatively short period of time of 2.4 ks at 578 K isothermally (Fig. 39) and of 42s at the range from 598 to 605 K during continuous heating at 0.17 K/s. Moreover, the unknown crystalline phase 1 has a wide area of homogeneity. Composition of crystalline particles determined by EDX analysis varied significantly ($Ge_{56-72}Al_{15-33}La_{12-17}$) whereas composition of amorphous phase ($Ge_{58.2}Al_{26.3}La_{15.6}$) was close to the alloy's composition. Thus, the transformation should be diffusive type. The thickness of the unknown phase 1 particle is about 50 nm on average while the size of the whole crystalline colony exceeds 10 μm. Simple calculation of the diffusion coefficient D using diffusion length l and time τ according to the relation $D \approx l^2/2\tau$ gives D values of about 1×10^{-18} and 3×10^{-17} m^2/s for single crystalline particles or 4×10^{-14} and 1×10^{-12} m^2/s for the whole colony at 578 and 598–605 K, respectively. Relatively high D values of alloying elements in the amorphous matrix obtained for the whole crystalline colony suggest an important role of grain boundary diffusion which can go much faster than the volume one.

The metastable nanocrystalline phase 2 was found to have a tetragonal lattice with lattice parameters of $a = 0.559$ nm and $c = 0.528$ nm [254]. According to EDX

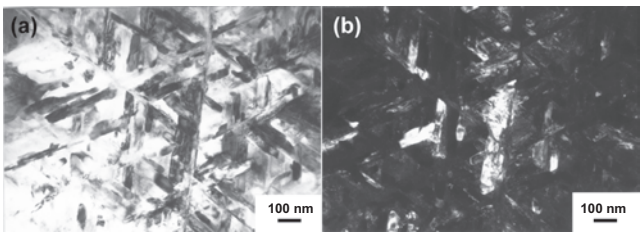


Figure 34. Structure of the $Ge_{60}Al_{25}La_{15}$ in as-solidified state, TEM: (a) bright-field and (b) dark-field images taken from the colony of nanocrystalline phase. Reprinted with permission from [252], D. V. Louzguine and A. Inoue, *Mater. Res. Bull.* 34, 1991 (1999). © 1999, Elsevier Science.

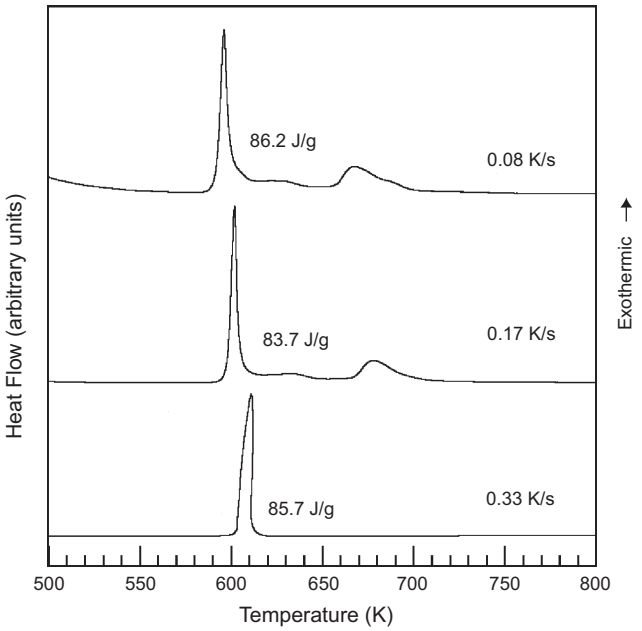


Figure 35. DSC curves of the $Ge_{60}Al_{25}La_{15}$ alloy taken at different heating rates. Reprinted with permission from [252], D. V. Louzguine and A. Inoue, *Mater. Res. Bull.* 34, 1991 (1999). © 1999, Elsevier Science.

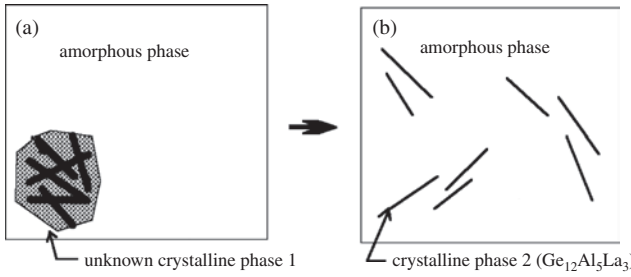


Figure 36. Structure changes in the $Ge_{60}Al_{25}La_{15}$ alloy scheme. (a) As-solidified state. (b) Heat treated for 2.4 ks at 578 K.

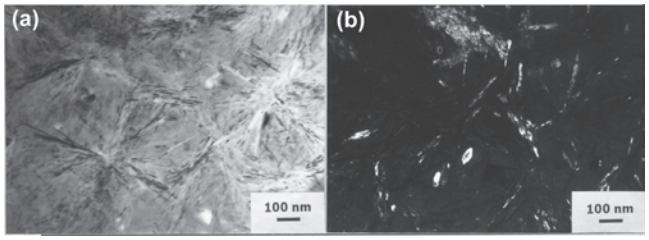


Figure 37. Structure of the $Ge_{60}Al_{25}La_{15}$ alloy heat treated for 2.4 ks at 578 K, TEM. (a) Bright-field and (b) Dark-field images. Reprinted with permission from [254], D. V. Louzguine and A. Inoue, *Mater. Sci. Eng. A* 304–306, 505 (2001). © 2001, Elsevier Science.

analysis (beam size 3 nm) compositions of three different crystals of crystalline phase 2 were found to be $Ge_{52}Al_{31}La_{17}$, $Ge_{64}Al_{20}La_{16}$, and $Ge_{59}Al_{25}La_{16}$ [254]. The average composition ($Ge_{58}Al_{26}La_{16}$) is very close to the alloy's composition and can be written as $Ge_{12}Al_5La_3$.

In order to study the details of the first phase transformation in the $Ge_{60}Al_{25}La_{15}$ alloy (i.e., amorphous phase + unknown crystalline phase 1 → unknown phase 2), an isothermal differential calorimetry was carried out at 573 and 578 K. Fraction transformed versus annealing time dependence is shown in Figure 39. The volume fraction transformed at the time t is assumed to scale with the fraction of the total heat released. The transformation kinetics can be described on the basis of Eq. (5) [255] for volume fraction (x) transformed as a function of time (t):

$$x(t) = 1 - \exp[-kt^n] \tag{5}$$

k is an effective rate constant and n is Avrami exponent [255]. This is a simplified case of Eq. (1). The Avrami plot of $\ln[-\ln(1-x)]$ vs $\ln(t)$ yields a straight line with slope n and intercept $n\ln(k)$. From the data shown in Figure 39 the average slope is calculated to be $n = 2.7$. The slope value slightly changes with temperature being 2.6 at 578 K and 2.8 at 573 K. Thus, the volume fraction transformed as a function of time can be described on the basis of Eq. (5)

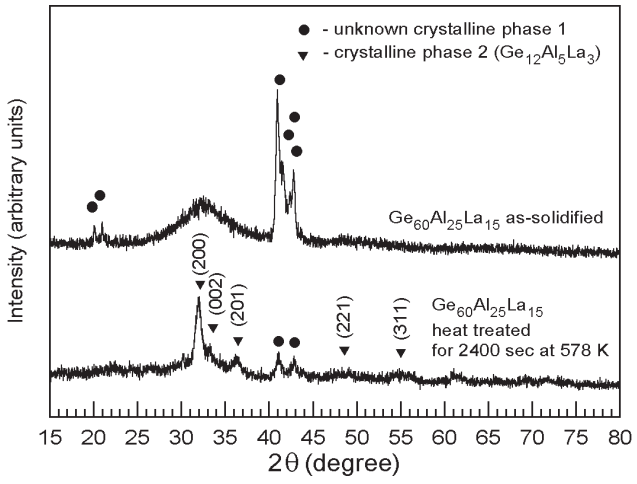


Figure 38. XRD patterns of the $Ge_{60}Al_{25}La_{15}$ alloy. Reprinted with permission from [254], D. V. Louzguine and A. Inoue, *Mater. Sci. Eng. A* 304–306, 505 (2001). © 2001, Elsevier Science.

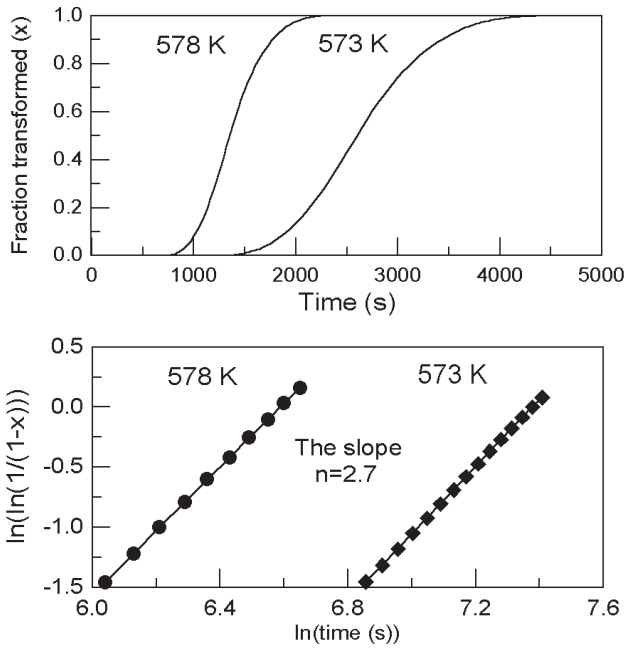


Figure 39. Volume fraction transformed versus annealing time dependence and Avrami plots for the $Ge_{60}Al_{25}La_{15}$ alloy constructed from the isothermal calorimetry measurements. Reprinted with permission from [252], D. V. Louzguine and A. Inoue, *Mater. Res. Bull.* 34, 1991 (1999). © 1999, Elsevier Science.

with $n = 2.7$. Such a value of the exponent n corresponds to accelerated nucleation and three-dimensional crystal growth by volume diffusion. The data show that the reaction is of diffusion-controlled type.

6.3. Complex Devitrification of Glassy Alloys

The $Ge_{68}Cr_{14}Al_{10}Ce_4Sm_4$ amorphous alloy was found to have a multistage crystallization behavior [256]. Four clearly separated heat effects marked as A, B, C, and D are observed in the DSC trace (see Fig. 4a). As has been mentioned the first prolonged heat effect A is related to the process of the short-range order change in the amorphous matrix leading to the segregation of the Ge atoms forming Ge-enriched zones prior to crystallization. However, the size of such zones is not larger than that of the medium-range order zones observed in the as-solidified state. As shown in Figure 40 no crystallinity is seen in the sample heat treated up to the completion of reaction A (see Fig. 4a).

The type of reaction B is the most interesting in the present alloy. In the sample heat treated in DSC up to 660 K reaction B is complete, but reaction C has not started yet (see Fig. 4a). Three broadened diffraction peaks belong to pure *Fd3m* Ge. As shown in Figure 41 the structure of the sample heat treated up to the completion of the reaction B at 0.083 K/s consists of the equiaxed nanoscale Ge particles with the size of about 3–20 nm homogeneously distributed in the amorphous matrix (Fig. 41a, b) and relatively large dendrites of an unknown multicomponent $Ge_{18}Cr_4AlCeSm$ crystalline phase of about 0.5 μm in size (Fig. 41d, e). Thus, precipitation of Ge nanoparticles as well as nucleation

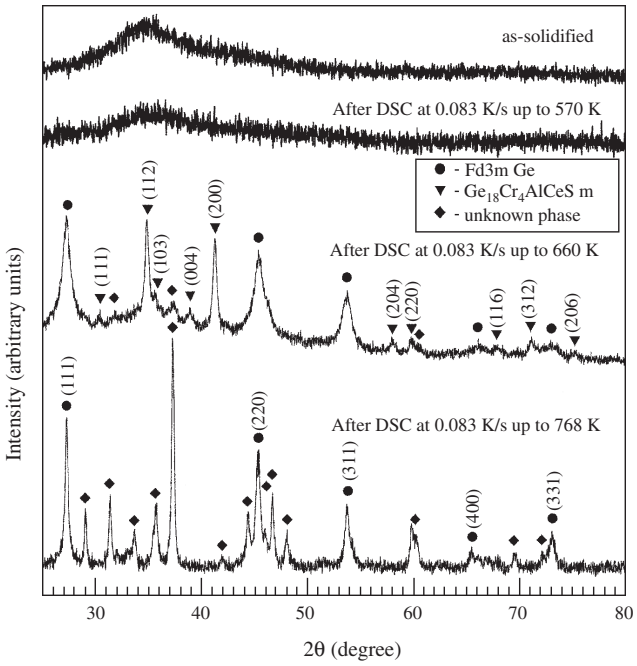


Figure 40. X-ray diffraction patterns of the $Ge_{68}Cr_{14}Al_{10}Ce_4Sm_4$ alloy in as-solidified and heat-treated using DSC state. Reprinted with permission from [256], D. V. Louzguine et al., *J. Mater. Sci.* 35, 5537 (2000). © 2000, Kluwer Academic.

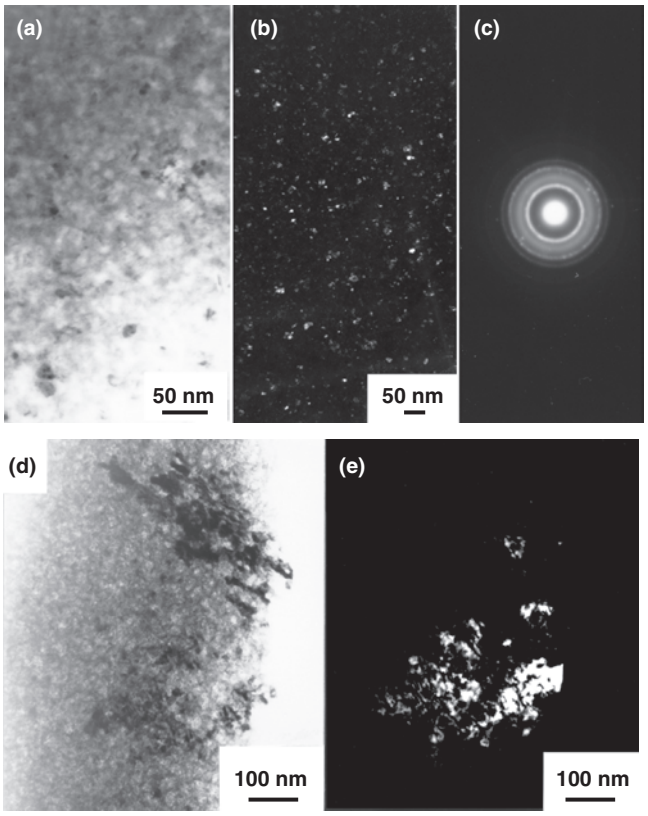


Figure 41. Microstructure of the $Ge_{68}Cr_{14}Al_{10}Ce_4Sm_4$ amorphous alloy. (a), (b), and (c) are bright-field and dark-field images and selected-area electron diffraction patterns, respectively, of the Ge nanoparticles embedded in an amorphous matrix. (d) and (e) are bright-field and dark-field images, respectively, of the $Ge_{18}Cr_4AlCeSm$ phase. Reprinted with permission from [256], D. V. Louzguine et al., *J. Mater. Sci.* 35, 5537 (2000). © 2000, Kluwer Academic.

and growth of the multicomponent $Ge_{18}Cr_4AlCeSm$ phase go simultaneously during a single heat effect B, although no eutectic-type coupled growth was observed. These two phases often have no common interface and are separated by the amorphous matrix (Fig. 41). Moreover, the size and volume fraction of the multicomponent phase depend on heating rate whereas the size and distribution of the Ge particles are independent of the rate. It can be suggested that Ge nanoparticles grew from the Ge—enriched zones formed during structural relaxation. As chemical composition of the $Ge_{18}Cr_4AlCeSm$ phase is close to composition of the amorphous matrix it likely nucleates by a polymorphous mechanism. $Ge_{18}Cr_4AlCeSm$ phase was found to have a tetragonal lattice of $a = 0.438$ nm and $c = 0.925$ nm [256]. This phase is a metastable one and disappears by reactions C and D (see Fig. 40 sample heat treated up to 768 K).

6.4. Solid-State Transformations Involving Quasicrystals

Mechanical alloying leads to formation of metastable and stable quasicrystalline phases in different alloy systems [257, 258]. For example, different structures including quasicrystals were obtained by heating a mechanically alloyed $Al_{65}Cu_{23}Fe_{12}$ powder mixture [259]. Solid-state reactions in mechanically alloyed nanocrystalline $Al_{65}Cu_{23}Fe_{12}$ powder were also studied [260]. A metastable disordered nanocrystalline phase forms as a result of mechanical alloying. Upon heating this phase undergoes a complex transformation, followed by partial decomposition and ordering. A metastable cubic phase transforms into an icosahedral quasicrystal in the temperature region 773–873 K [260].

In Al-based alloys an ordering transformation of the icosahedral phase at room temperature was observed. A transformation from higher symmetry of quasicrystals results in several variants of the product phase; for example, six variants of the decagonal phase are produced on transformation from the icosahedral phase [261]. The transformation of quasicrystalline phases to crystals involves the production of a high density of planar faults which produce streaking of reciprocal spots. The quasicrystals may transform into several intermediate metastable phases before transforming into the stable crystalline phase.

7. APPLICATIONS

Metallic nanomaterials are mostly used for structural and magnetic applications.

There are some reasons to believe that the problem of low ductility of nanostructured metallic materials subjected for structural applications will be solved in the near future and the production cost factor will be the critical one limiting their application. Nevertheless, nanostructured Ti could be used to replace commercial Ti-6Al-4V alloy in biomedical applications as an implantate. High mechanical properties were recently obtained in nanostructured Ti obtained by

severe plastic deformation [262]. Nanostructured Ti is also suitable for use in microthreaded connections.

The nanocrystalline Al-based alloys exhibiting high tensile strength of about 1 GPa have been already commercialized under the commercial name GIGAS. The GIGAS alloys also have higher specific strength and higher specific Young's modulus which are superior to those for conventional Al-based and Ti-based alloys. These advantages are already used in materials for some application fields such as rapidly repeating machinery parts requiring high specific strength and high specific modulus, main construction parts in robots, and high elevated temperature die casting materials, etc. [167].

The martensitic 9Cr-ODS (of Fe–9Cr–0.13C–2W–0.2Ti–$0.35Y_2O_3$ (mass%) composition) and ferritic 12Cr-ODS steels reinforced with stable Y_2O_3 oxide particles on a nanoscale are promising fuel cladding materials in fast nuclear reactors [263]. Oxide particles serve as blocks for mobile dislocations and as sinks for radiation defects at the particle–matrix interfaces resulting in high values of high-temperature strength and ductility. The alloys are produced by powder metallurgy and consolidation by hot isostatic pressure or hot extrusion.

Finemet $Fe_{73.5}Cu_1Nb_3Si_{13.5}B_9$, Nanoperm $Fe_{84}Zr_{3.5}Nb_{3.5}B_8Cu_1$, and Hitperm $Fe_{44}Co_{44}Zr_7B_4Cu_1$ alloys are used as soft magnetic materials. The feature limiting their application is low critical thickness.

Planar micromotors have the potential for high-speed control applications. These systems require closed loop control and therefore involve not only the motor but also associated high-speed control circuitry. Magnetic micromotors have been constructed by using deep X-ray lithography and metal plating. The devices consist of a stator that is rigidly attached to the substrate and a rotor that is fabricated as a fully released, free part made of Ni of 100 μm in thickness. The rotor is assembled onto the stator shaft with submicrometer tolerances. The structure is driven by an external rotating magnet or a fixed electromagnet. It is reported that rotational speeds of up to 8000 rpm have been obtained and maintained for several days [264]. Modulating the spectral properties of synchrotron radiation, different regimes of X-ray lithography lead to the mass fabrication of nanostructure as high resolution X-ray lithography or classical X-ray lithography is able to obtain a nanometer resolution below 50 nm [265].

Materials with reduced dimensions in the form of isolated nanoparticles, thin wires, or thin films are used in catalysts and semiconductor devices utilizing single or multilayer quantum structures.

Nanostructured surfaces with enhanced corrosion resistance, hardness, and wear resistance can be used as protective coatings in which the properties of a thin surface layer are improved by means of creating a nanoscale microstructure in a thin surface region.

Materials in the form of an array of quantum dots connected by nanowires synthesized by nanolithography [266], the tip of a tunneling microscope, or electron or ion beams are expected to play a key role in the production of the next generation of electronic devices such as terrabit memories [267], quantum computers, etc.

Nanostructured materials offer potential applications for hydrogen storage materials. High absorption–desorption rates can be achieved by decreasing the grain size [268].

High localized corrosion resistance of nanocrystalline Fe–18%Cr–8%Ni stainless steel prepared by sputter deposition [269] offers another possibility of usage as a corrosion resistant material.

In conclusion one must say that although in our days applications of metallic nanomaterials, especially as structural materials, are still very limited without any doubt nanostructured materials and nanocomposites will play a leading role in the materials science of the present century.

GLOSSARY

3DAP Three-dimensional atom probe technique.

Bs Saturation magnetic flux density.

Devitrification Solid-state reaction related to the formation of a phase with a long-range order from a glassy (amorphous) phase with a short-range order. Can be also called "crystallization" but "devitrification" is more precise term as there is the possibility of confusion with crystallization reaction from the liquid. Process opposite to vitrification (glass formation).

Hc Coercivity.

hcp Hexagonal close packed lattice.

MRO Medium-range order zones, typically 0.5–2 nm in size or a few interatomic distances. An intermediate ordered state between glassy (disordered on a long-range-scale) and long-range ordered crystalline state. MRO is characterized by a diffuse boundaries between ordered regions.

n Avrami exponent.

SAED Selected-area electron diffraction pattern. Taken via TEM of the local area.

UTS Ultimate tensile strength, close on meaning to σ_f.

REFERENCES

1. A. Wilm, *Metallurgie* 8, 225 (1911).
2. A. Guinier, *Nature* 142, 569 (1938); G. D. Preston, *Nature* 142, 570 (1938).
3. K. Ohshima, S. Yatsuya, and J. Harada, *J. Phys. Soc. Jpn.* 50, 3071 (1981).
4. J. Harada, S. Yao, and A. Ichimiya, *J. Phys. Soc. Jpn.* 48, 1625 (1980).
5. H. Gleiter, *Progr. Mater. Sci.* 33, 223 (1989).
6. "Nanomaterials: Synthesis, Properties, and Applications" (A. S. Edelstein and R. C. Cammarata, Eds.). Institute of Physics, Bristol, 1996.
7. P. G. Le Comber, A. Madan, and W. E. Spear, "Electronic and Structural Properties of Amorphous Semiconductors" (P. G. Le Comber and J. Mort, Eds.), p. 373. Academic Press, London/New York, 1973.
8. B. Cantor, *Mater. Sci. Forum* 307, 143 (1999).
9. Ebrahimi, G. R. Bourne, M. S. Kelly, and T. E. Matthews, *Nanostruct. Mater.* 11, 343 (1999).
10. H. J. Fecht, *Nanostructured Mater*. 6, 33 (1995).
11. R. Z. Valiev, *Mater. Sci. Eng. A* 234–236, 59 (1997).
12. V. M. Segal, *Mater. Sci. Eng. A* 197, 157 (1995).
13. A. L. Greer, *Science* 267, 1947 (1995).

14. P. E. Bradley, B. Janossy, and D. Landolt, *J. Appl. Electrochem*. 31, 137 (2001).
15. Yu. A. Babanov, I. V. Golovshchikova, F. Boscherini, and S. Mobilio, *Nanostruct. Mater.* 6, 601 (1995).
16. H. J. Fecht, E. Hellstern, Z. Fu, and W. L. Johnson, *Metall. Trans. A* 21, 2333 (1990).
17. H. J. Fecht, *Nanostruct. Mater.* 6, 33 (1995).
18. D. Oleszak and P. H. Shingu, *J. Appl. Phys.* 79, 2975 (1996).
19. J. Eckert, J. C. Holzer, C. E. Krill, and W. L. Johnson, *J. Mater. Res.* 7, 1751 (1992).
20. C. C. Koch, *Nanostruct. Mater.* 2, 109 (1993).
21. A. W. Weeber and H. Bakker, *Physica B* 153, 93 (1988).
22. R. K. Islamgaliev, R. Kuzel, S. N. Mikov, A. V. Igo, J. Burianek, F. Chmelik, and R. Z. Valiev, *Mater. Sci. Eng. A* 266, 205 (1999).
23. A. A. Nazarov, A. E. Romanov, and R. Z. Valiev, *Scripta Mater.* 34, 729 (1996).
24. M. Murayama, J. M. Howe, H. Hidaka, and S. Takaki, *Science* 295, 2433 (2002).
25. A. Inoue, C. Fan, J. Saida, and T. Zhang, *Sci. Technol. Adv. Mater.* 1, 73 (2000).
26. J. Saida, M. Matsushita, and A. Inoue, *Appl. Phys. Lett.* 77, 1287 (2000).
27. A. R. Yavari and D. Negri, *Nanostruct. Mater.* 8, 969 (1997).
28. L. Battezzati, C. Antonione, G. Riontino, F. Marino, and H. R. Sinning, *Acta Metall. Mater.* 39, 2107 (1991).
29. W. N. Myung, L. Battezzati, M. Baricco, K. Oaki, A. Inoue, and T. Masumoto, *Mater. Sci. Eng. A* 179 & 180, 371 (1994).
30. U. Koster and J. Meinhardt, *Mater. Sci. Eng. A* 178, 271 (1994).
31. J. W. Cahn, *Acta Metall*. 9, 795 (1961).
32. S. Schneider, P. Thiyagarajan, U. Geyer, and W. L. Johnson, *Physica B* 241–243, 918 (1997).
33. A. Inoue, K. Ohtera, A. P. Tsai, and T. Masumoto, *Jpn. J. Appl. Phys.* 27, L280 (1988).
34. A. Inoue, K. Ohtera, and T. Masumoto, *Jpn. J. Appl. Phys.* 27, L1796 (1988).
35. Y. He, S. J. Poon, and G. J. Shiflet, *Science* 241, 1640 (1988).
36. A. Inoue, K. Ohtera, A. P. Tsai, and T. Masumoto, *Jpn. J. Appl. Phys.* 27, L736 (1988).
37. A. Inoue, T. Zhang, K. Kita, and T. Masumoto, *Mater. Trans. JIM* 30, 870 (1989).
38. S. R. Elliot, "Physics of Amorphous Materials." Longman Group, Harlow, 1990.
39. A. Inoue and H. M. Kimura, *J. Metastable Nanocrystalline Mater.* 9, 41 (2001).
40. A. Inoue, K. Ohtera, A. P. Tsai, and T. Masumoto, *Jpn. J. Appl. Phys.* 27, L479 (1988).
41. G. J. Shihlet, Y. He, and S. J. Poon, *J. Appl. Phys.* 64, 6863 (1988).
42. A. Inoue, N. Matsumoto, and T. Masumoto, *Mater. Trans. JIM* 31, 493 (1990).
43. D. R. Allen, J. C. Foley, and J. H. Perepezko, *Acta Mater.* 46, 431 (1998).
44. A. Inoue and H. M. Kimura, *J. Light Met.* 1, 31 (2001).
45. Y. H. Kim, A. Inoue, and T. Masumoto, *Mater. Trans. JIM* 32, 331 (1991).
46. M. Gogebakan, P. J. Warren, and B. Cantor, *Mater. Sci. Eng. A* 226–228, 168 (1997).
47. J. C. Foley, D. R. Allen, and J. H. Perepezko, *Scripta Mater.* 35, 655 (1996).
48. M. Calin, A. Rudiger, and U. Koester, *J. Metastable Nanocryst. Mater.* 8, 359 (2000).
49. K. Hono, Y. Zhang, A. P. Tsai, A. Inoue, and T. Sakurai, *Scripta Mater.* 32, 191 (1995).
50. T. B. Massalski, "Binary Alloy Phase Diagrams." ASM International, Materials Park, OH, 1990.
51. A. R. Yavari and O. Drbohlav, *Mater. Trans. JIM* 36, 896 (1995).
52. A. R. Yavari and D. Negri, *Nanostruct. Mater.* 8, 969 (1997).
53. K. Pekala, P. Jaskiewicz, J. Latuch, and A. Kokoszkiewicz, *J. Non-Cryst. Solids* 211, 72 (1997).
54. N. Bassim, C. S. Kiminami, and M. J. Kaufman, *J. Non-Cryst. Solids* 273, 271 (2000).
55. D. V. Louzguine and A. Inoue, *J. Non-Cryst. Solids* 311, 281 (2002).
56. D. V. Louzguine and A. Inoue, *Appl. Phys. Lett.* 78, 3061 (2001).
57. K. F. Kelton, T. K. Croat, A. K. Gangopadhyay, L. Q. Xing, A. L. Greer, M. Weyland, X. Lic, and K. Rajan, *J. Non-Cryst. Solids*, in press.
58. D. V. Louzguine and A. Inoue, *Scripta Mater.* 43, 371 (2000).
59. A. N. Kolmogorov, *Isz. Akad. Nauk. USSR Ser. Matem.* 3, 355 (1937).
60. M. W. A. Johnson and K. F. Mehl, *Trans. Am. Inst. Mining. Met. Eng.* 135, 416 (1939).
61. M. Avrami, *J. Chem. Phys.* 9, 177 (1941).
62. K. F. Kelton, *J. Non-Cryst. Solids* 163, 283 (1993).
63. L. Q. Xing, J. Eckert, W. Loser, L. Schultz, and D. M. Herlach, *Philos. Mag. A* 79, 1095 (1999).
64. U. Koster, J. Meinhardt, S. Roos, and H. Liebertz, *Appl. Phys. Lett.* 69, 179 (1996).
65. L. Q. Xing, J. Eckert, W. Loser, and L. Schultz, *Appl. Phys. Lett.* 73, 2110 (1998).
66. M. W. Chen, T. Zhang, A. Inoue, A. Sakai, and T. Sakurai, *Appl. Phys. Lett.* 75, 1697 (1999).
67. A. Inoue, T. Zhang, J. Saida, M. Matsushita, M. W. Chen, and T. Sakurai, *Mater. Trans. JIM* 40, 1137 (1999).
68. J. Saida, M. Matsushita, C. Li, and A. Inoue, *Appl. Phys. Lett.* 76, 3558 (2000).
69. L. Wang, L. Ma, and A. Inoue, *Mater. Trans. JIM* 43, 2346 (2002).
70. D. V. Louzguine and A. Inoue, *Appl. Phys. Lett.* 78, 1841 (2001).
71. J. Saida and A. Inoue, *J. Non-Cryst. Solids* 312–314, 502 (2002).
72. J. Saida, M. Matsushita, and A. Inoue, *J. Appl. Phys.* 88, 6081 (2000).
73. B. S. Murty, D. H. Ping, M. Ohnuma, and K. Hono, *Acta Mater.* 49, 3453 (2001).
74. J. Saida, M. Matsushita, and A. Inoue, *J. Alloys Compounds* 342, 18 (2002).
75. J. Saida, M. Matsushita, and A. Inoue, *Intermetallics* 10, 1089 (2002).
76. J. Saida, M. Matsushita, and A. Inoue, *J. Non-Cryst. Solids* 312, 617 (2002).
77. J. Z. Jiang, A. R. Rasmussen, C. H. Jensen, Y. Lin, and P. L. Hansen, *Appl. Phys. Lett.* 80, 2090 (2002).
78. E. Matsubara, M. Sakurai, T. Nakamura, M. Imafuku, S. Sato, J. Saida, and A. Inoue, *Scripta Mater.* 44, 2297 (2001).
79. A. Inoue, H. M. Kimura, and K. Kita, "New Horizons in Quasicrystals" (A. I. Goldman, D. J. Sordelet, P. A. Thiel, and J. M. Dubois, Eds.). World Scientific, Singapore, 1997.
80. D. V. Louzguine, A. Takeuchi, and A. Inoue, *J. Mater. Sci.* 35, 5537 (2000).
81. D. V. Louzguine, A. Inoue, M. Saito, and Y. Waseda, *Scripta Mater.* 42, 289 (2000).
82. D. V. Louzguine and A. Inoue, *Mater. Trans. JIM* 40, 485 (1999).
83. D. V. Louzguine and A. Inoue, *Mater. Trans. JIM* 39, 245 (1998).
84. D. V. Louzguine and A. Inoue, in "Amorphous and Nanocrystalline Materials, Preparation, Properties and Application" (A. Inoue and K. Hashimoto, Eds.), Advances in Materials Research, Vol. 3. Springer-Verlag, Berlin/Heidelberg, 2001.
85. A. M. Morales and C. M. Lieber, *Science* 279, 208 (1998).
86. R. Ferre, K. Ounadjela, J. M. George, L. Piraux, and S. Dubois, *Phys. Rev. B* 56, 14066 (1997).
87. S. Iijima, *Nature* 354, 56 (1991).
88. N. G. Chopra, R. J. Luyken, K. Cherrey, V. H. Crespi, M. L. Cohen, S. G. Loiue, and A. Zettl, *Science* 269, 966 (1995).
89. J. H. Weaver, *J. Phys. Chem. Solids* 53, 1433 (1992).

90. B. Cantor, C. M. Allen, R. Dunin-Burkowski, M. H. Green, J. L. Hutchinson, K. A. Q. O'Reilly, A. K. Petford-Long, P. Schumacher, J. Sloan, and P. J. Warren, *Scripta Mater.* 44, 2055 (2001).
91. H. Gleiter, *Nanostruct. Mater.* 6, 3 (1995).
92. G. Decher, *Science* 277, 5330 (1997).
93. Y. Lvov, G. Decher, and G. Sukhorukov, *Macromolecules* 26, 5396 (1993).
94. J. B. Wachtman and Z. H. Kalman, "Characterization of Materials." Butterworth–Heinenman, Greenwich, 1993.
95. A. R. Yavari, A. Le Mouleca, A. Inoue, W. J. Botta, F. G. Vaughan, and A. Kvick, *Mater. Sci. Eng. A* 304–306, 34 (2001).
96. P. Scherrer, *Nachr. Gottingen* 2, 98 (1918).
97. B. E. Warren and B. L. Averbach, *J. Appl. Phys.* 21, 595 (1950).
98. G. K. Williamson and W. H. Hall, *Acta Metall.* 1, 22 (1953).
99. H. H. Tian and M. Atzmon, *Philos. Mag. A* 79, 1769 (1999).
100. I. V. Alexandrov, K. Zhang, and K. Lu, *Ann. Chim.* 21, 407 (1996).
101. K. Zhang, I. V. Alexandrov, R. Z. Valiev, and K. Lu, *J. Appl. Phys.* 88, 5617 (1996).
102. R. W. Siegel, *J. Phys. Chem. Solids* 55, 1097 (1994).
103. G. Le Caer and P. Delcroix, *Nanostruct. Mater.* 7, 127 (1996).
104. J. Löffler, J. Weissmüller, and H. Gleiter, *Nanostruct. Mater.* 6, 567 (1994).
105. G. J. Thomas, R. W. Siegel, and J. A. Eastman, *Scr. Metall. Mater.* 24, 201 (1990).
106. R. W. Siegel and G. J. Thomas, *Ultramicroscopy* 40, 376 (1992).
107. E. A. Stern, R. W. Siegel, M. Newville, P. G. Sanders, and D. Haskel, *Phys. Rev. Lett.* 75, 3874 (1995).
108. R. Z. Valiev, V. Yu. Gertsman, and O. A. Kaibyshev, *Phys. Status Solidi* 77, 97 (1983).
109. R. Z. Valiev and O. A. Kaibyshev, *J. Phys.* 46, 641 (1985).
110. P. H. Pumphrey, in "Grain Boundary Structure and Properties" (G. A. Chadwick and D. A. Smith, Eds.). Academic Press, UK, 1976.
111. D. Shechtman, L. A. Blech, D. Gratias, and J. W. Cahn, *Phys. Rev. Lett.* 53, 1951 (1984).
112. A. Inoue, H. M. Kimura, and T. Masumoto, *J. Mater. Sci.* 22, 1758 (1987).
113. K. V. Rao, J. Fildler, and H. S. Chen, *Europhys. Lett.* 1, 647 (1986).
114. A. Inoue, L. Arnberg, B. Lehtinen, M. Oguchi, and T. Msumoto, *Metall. Trans.* 17A, 1657 (1986).
115. Y. Elser and C. L. Henley, *Phys. Rev. Lett.* 55, 2883 (1985).
116. A. Inoue, T. Zhang, S. Ishihara, J. Saida, and M. Matsushita, *Scripta Mater.* 44, 1615 (2001).
117. A. K. Srivastava and S. Ranganathan, *Acta Mater*. 44, 2935 (1996).
118. A. Inoue and H. M. Kimura, *J. Metastable Nanocryst. Mater.* 41 (2001).
119. A. Inoue, C. Fan, J. Saida, and T. Zhang, *Sci. Technol. Adv. Mater.* 1, 73 (2000).
120. T. Wada, Master's Thesis, Tohoku University, 2003.
121. R. Z. Valiev and O. A. Kaibyshev, *J. Phys.* 46, C4-641 (1985).
122. J. Lian, R. Z. Valiev, and B. Baudelet, *Acta Netall. Mater.* 43, 4165 (1995).
123. J. R. Groza and R. J. Dowding, *Nanostruct. Mater.* 7, 749 (1996).
124. H. J. Hofler and R. S. Averback, *Scripta Metall. Mater.* 24, 2401 (1990).
125. S. K. Ganapathi, D. M. Owen, and A. H. Chokshi, *Scripta Metall. Mater.* 25, 2699 (1991).
126. T. R. Malow and C. C. Koch, *Acta Mater.* 45, 2177 (1997).
127. P. A. Beck, J. C. Kremer, L. J. Demer, and M. L. Holzworth, *Trans. AIME* 175, 372 (1948).
128. J. E. Burke, *Trans. AIME* 180, 73 (1949).
129. S. K. Ganapathi, D. M. Owen, and A. H. Chokshi, *Scripta Metall. Mater.* 25, 2699 (1991).
130. J. Z. Jiang, *Nanostruct. Mater.* 9, 245 (1997).
131. Kh. Ya. Mulyukov, G. F. Korznikova, R. Z. Abdulov, and R. Z. Valiev, *J. Magn. Magn. Mater.* 90, 207 (1990).
132. D. V. Louzguine, M. S. Ko, S. Ranganathan, and A. Inoue, in "Proceedings of The Fourth Pacific Rim International Conference on Advanced Materials and Processing (PRICM4)," 11–15 December 2001, Honolulu, HI (S. Hanada, Z. Zhong, S. W. Nam, and R. N. Wright, Eds.), p. 67. Japan Institute of Metals, 2001.
133. P. Knauth, A. Charai, and P. Gas, *Scripta Metall. Mater.* 28, 325 (1993).
134. J. Weissmuller, *J. Mater. Res.* 9, 4 (1994).
135. R. Z. Valiev, G. F. Korznikova, Kh. Ya. Mulyukov, R. S. Mishra, and A. K. Mukherjee. *Philoso. Mag. B* 75, 803 (1997).
136. P. G. Sanders, J. A. Eastman, and J. R. Weertman, *Acta Mater.* 45, 4019 (1997).
137. C. C. Koch, D. G. Morris, K. Lu, and A. Inoue, *MRS Bull.* 2, 54 (1999).
138. N. Wang, Z. Wang, K. T. Aust, and U. Erb, *Mater. Sci. Eng. A* 237, 150 (1997).
139. A. Robertson, U. Erb, and G. Palumbo, *Nanostruct. Mater.* 12, 1035 (1999).
140. C. C. Koch, *J. Metastable Nanocrystalline Mater*. 18, 9 (2003).
141. A. A. Karimpoor, U. Erb, K. T. Aust, Z. Wang, and G. Palumbo, *J. Metastable Nanocrystalline Mater.* 13, 415 (2001).
142. G. W. Nieman, J. W. Weertman, and R. W. Siegel, *Scripta Metall. Mater.* 24, 145 (1990).
143. A. H. Chokshi, A. Rosen, J. Karch, and H. Gleiter, *Scripta Metall.* 23, 1679 (1989).
144. U. Erb, *Nanostruct. Mater.* 6, 533 (1995).
145. P. G. Sanders, C. J. Youngdahl, and J. R. Weertman, *Mater. Sci. Eng. A* 234–236, 77 (1997).
146. T. G. Nieh and J. Wadsworth, *Scripta Metall. Mater.* 25, 955 (1991).
147. R. A. Masumura, P. M. Hazzledine, and C. S. Pande, *Acta Mater.* 46, 4527 (1998).
148. H. Van Swygenhoven, M. Spaczer, and A. Caro, *Acta Mater.* 47, 3117 (1999).
149. C. J. Youngdahl, J. R. Weertman, R. C. Hugo, and H. H. Kung, *Scripta Mater*. 44, 1475 (2001).
150. R. O. Scattergood and C. C. Koch, *Scripta Metall. Mater.* 27, 1195 (1992).
151. J. Lian, B. Baudelet, and A. A. Nazarov, *Mater. Sci. Eng. A* 172, 23 (1993).
152. V. G. Gryaznov, M. Yu. Gutkin, A. E. Romanov, and L. I. Trusov, *J. Mater. Sci.* 38, 4359 (1993).
153. R. A. Masumura, P. M. Hazzledine, and C. S. Pande, *Acta Mater.* 46, 4527 (1998).
154. P. G. Sanders, M. Rittner, E. Kiedaisch, J. R. Weertman, H. Kung, and Y. C. Lu, *Nanostruct. Mater.* 9, 433 (1997).
155. B. Cai, Q. P. Kong, L. Lu, and K. Lu, *Mater. Sci. Eng. A* 286, 188 (2000).
156. A. B. Witney, P. G. Sanders, J. R. Weertman, and J. A. Eastman, *Scipta Met. Mater.* 33, 2025 (1995).
157. A. Vinogradov, Y. Kaneko, K. Kitagawa, S. Hashimoto, V. Stolyarov, and R. Valiev, *Scripta Mater.* 36, 1345 (1997).
158. S. R. Agnew and J. R. Weertman, *Mater. Sci. Eng. A* 244, 145 (1998).
159. K. Ohtera, A. Inoue, T. Terabayashi, H. Nagahama, and T. Masumoto, *Mater. Trans. JIM* 33, 775 (1992).
160. A. K. Vasudevan and R. O. Doherty, "Aluminum Alloys." Academic Press, London, 1989.
161. J. W. Martin, "Micromechanism in Particle Hardened Alloys." Cambridge Univ. Press, Cambridge, UK, 1980.
162. R. O. England, J. R. Pickens, K. S. Kumar, and T. J. Langan, "Dispersion Strengthened Aluminum Alloys" (Y. W. Kim and W. M. Griffth, Eds.). The Minerals, Metal & Materials Society, 1988.
163. V. R. Bennett, W. D. Nix, and A. S. Tettleman, "The Principles of Engineering Materials." Prentice-Hall, Englewood Cliffs, NJ, 1973.
164. R. F. Decker, *Metall. Trans.* 11, 2495 (1973).

165. J. A. Wert, "Strength of Metal Alloys" (R. C. Gifkins, Ed.). Pergamon Press, Oxford, 1980.
166. Y. W. Kim and W. M. Griffith, "PM Aerospace Materials." Shrewdburg, UK, 1984.
167. A. Inoue and H. Kimura, *J. Light Metals* 1, 31 (2001).
168. A. Kato, A. Inoue, H. Horikiri, and T. Masumoto, *Mater. Trans. JIM* 36, 977 (1995).
169. A. Inoue, in "Handbook on the Physics and Chemistry of Rare Earths" (K. Gschneidner and L. Eyring, Eds.), Vol. 24. Elsevier/North-Holland, 1997.
170. H. Horikiri, A. Kato, A. Inoue, and T. Masumoto, *Mater. Sci. Eng. A* 179/180, 702 (1994).
171. "Handbook of Advanced Magnesium Technology." Kallos, Tokyo, 2000.
172. A. Kato, Doctor Thesis, Tohoku University, 1995.
173. Y. Kawamura, K. Hayashi, A. Inoue, and T. Masumoto, *Mater. Trans.* 42, 1172 (2001).
174. A. Inoue, "Bulk Amorphous Alloys." Trans Tech, Zurich, 1998.
175. A. Inoue, Y. Kawamura, M. Matsushita, K. Hayashi, and J. Koike, *J. Mater. Res.* 16, 1894 (2001).
176. "Metal Databook." Japan Inst. Metals, Maruzen, Tokyo, 1983.
177. E. Abe, Y. Kawamura, K. Hayashi, and A. Inoue, *Acta Mater.* 50, 3845 (2002).
178. P. Villars, A. Prince, L. Zabdyr, and Z. Moser, "Handbook of Ternary Alloy Phase Diagrams." ASM International, Metals Park, OH, 1995.
179. T. Tei-Ohkawa, S. Takeuchi, and K. Edagawa, *J. Non-Cryst. Solids* 189, 25 (1995).
180. Y. Yokoyama, A. P. Tsai, A. Inoue, and T. Masumoto, *Mater. Trans. JIM* 32, 1089 (1991).
181. Y. Yokoyama, A. Inoue, and T. Masumoto, *Mater. Trans. JIM* 34, 135 (1993).
182. A. Inoue, H. M. Kimura, K. Sasamori, and T. Masumoto, *Mater. Trans. JIM* 35, 85 (1994).
183. A. Inoue, H. M. Kimura, K. Sasamori, and T. Masumoto, *Mater. Trans. JIM* 36, 6 (1995).
184. A. Inoue, H. M. Kimura, M. Watanabe, and A. Kawabata, *Mater. Trans. JIM* 38, 756 (1997).
185. A. Inoue and H. M. Kimura, *Mater. Sci. Forum* 235–238, 873 (1997).
186. A. Inoue, H. M. Kimura, and K. Sasamori, "Chemistry and Physics of Nanostructures and Related Non-Equilibrium Materials" (E. Ma, B. Fultz, R. Shull, J. Morral, and P. Nash, Eds.). The Minerals, Metals & Materials Society, 1997.
187. A. Inoue, H. M. Kimura, K. Sasamori, and K. Kita, *Aluminum Alloys* 3, 1841 (1998).
188. A. Inoue, H. M. Kimura, and K. Sasamori, *Adv. Mater.* 4, 91 (1998).
189. A. K. Vasudevan and R. O. Doherty, "Aluminum Alloys." Academic Press, London, 1989.
190. D. H. Bae, M. H. Lee, K. T. Kim, W. T. Kim, and D. H. Kim, *J. Alloys Compounds* 342, 445 (2002).
191. D. H. Bae, S. H. Kim, W. T. Kim, and D. H. Kim, *Mater. Trans. JIM* 42, 2144 (2001).
192. A. Inoue, K. Ohtera, A. P. Tsai, and T. Masumoto, *Jpn. J. Appl. Phys.* 27, L280 (1988).
193. Y. H. Kim, A. Inoue, and T. Masumoto, *Mater. Trans. JIM* 31, 747 (1990).
194. Y. H. Kim, A. Inoue, and T. Masumoto, *Mater. Trans. JIM* 32, 599 (1991).
195. J. H. Perepezko, R. J. Hebert, and W. S. Tong, *Intermetallics* 10, 1079 (2002).
196. A. Inoue and H. M. Kimura, *Mater. Sci. Forum* 235–238, 873 (1997).
197. A. Inoue, K. Nakazato, Y. Kawamura, A. P. Tsai, and T. Masumoto, *Mater. Trans. JIM* 32, 331 (1991).
198. A. Inoue, K. Nakazato, Y. Kawamura, A. P. Tsai, and T. Msumoto, *Mater. Trans. JIM* 35, 95 (1994).
199. Y.-H. Kim, A. Inoue, and T. Masumoto, *Mater. Trans. JIM* 31, 747 (1990).
200. A. Inoue, Y. H. Kim, and T. Masumoto, *Mater. Trans. JIM* 33, 487 (1992).
201. A. Inoue, H. M. Kimura, and K. Kita, "New Horizons in Quasicrystals" (A. I. Goldman, D. J. Sordelet, P. A. Thiel, and J. M. Dubois, Eds.), p. 256. World Scientific, Singapore, 1997.
202. Z. C. Zhong, X. Y. Jiang, and A. L. Greer, *Mater. Sci. Eng. A* 226–228, 531 (1997).
203. A. L. Greer, *Mater. Sci. Eng. A* 304–306, 68 (2001).
204. H. S. Kim and S. I. Hong, *Acta Mater.* 47, 2059 (1999).
205. K. Ohtera, A. Inoue, T. Terabayashi, H. Nagahama, and T. Masumoto, *Mater. Trans. JIM* 33, 775 (1992).
206. D. V. Louzguine and A. Inoue, *J. Mater. Res.* 17, 1014 (2002).
207. A. Takeuchi and A. Inoue, *Mater. Trans. JIM* 41, 1372 (2000).
208. Y. H. Kim, A. Inoue, and T. Masumoto, *Mater. Trans. JIM* 32, 599 (1991).
209. H. Chen, Y. He, G. J. Shiflet, and S. J. Poon, *Scripta Mater*. 25, 1421 (1991).
210. A. Inoue, T. Zhang, M. W. Chen, T. Sakurai, J. Saida, and M. Matsushita, *Appl. Phys. Lett.* 76, 967 (2000).
211. S. Takeuchi, "Quasicrystals." Sangyotosho, Tokyo, 1992.
212. K. Fujita, A. Inoue, and T. Zhang, *Mater. Trans. JIM* 41, 1448 (2000).
213. M. Heilmaier, *J. Mater. Processing Technol.* 117, 374 (2001).
214. C. Ma and A. Inoue, *Mater. Trans.* 44, 188 (2003).
215. H. Kato and A. Inoue, *Mater. Trans. JIM* 38, 793 (1997).
216. H. Kato, T. Hirano, A. Matsuo, Y. Kawamura, and A. Inoue, *Scripta Mater.* 43, 503 (2000).
217. O. A. Kaibyshev, "Superplasticity of Alloys, Intermetallides, and Ceramics," p. 317. Springer-Verlag, Berlin/New York, 1992.
218. K. Higashi, T. Mukai, S. Tanimura, A. Inoue, T. Masumoto, K. Kita, K. Ohtera, and J. Nagahora, *Scr. Metall.* 26, 191 (1992).
219. J. Mukhopadhyay, G. Kaschner, and A. K. Mukherjee, *Scripta Metall. Mater.* 24, 857 (1990).
220. A. A. Fedorov, M. Yu. Gutkin, and I. A. Ovid'ko, *Acta Mater*. 51, 887 (2003).
221. G. Fanta, R. Bohn, M. Dahms, T. Klassen, and R. Bormann, *Intermetallics* 9, 45 (2001).
222. R. Bohn, G. Fanta, T. Klassen, and R. Bormann, *Scripta Mater.* 44, 1479 (2001).
223. R. S. Mishra, R. Z. Valiev, S. X. McFadden, R. K. Islamgaliev, and A. Mukherjee, *Philoso. Mag. A* 81, 37 (2001).
224. X. McFadden, R. S. Mishra, R. Z. Valiev, A. P. Zhilyaev, and A. K. Mukherjee, *Nature* 6729, 684 (1999).
225. A. P. Zhilyaev, *Mater. Phys. Mech.* 1, 98 (2000).
226. H. M. Kimura, A. Inoue, K. Sasamori, and Y. Kawamura, *J. Jpn. Inst. Light Met.* 47, 487 (1997).
227. A. Inoue, Y. Kawamura, M. Matsushita, and K. Hayashi, *Mater. Sci. Forum* 386–388, 509 (2002).
228. D. H. Bae, H. K. Lim, S. H. Kim, D. H. Kim, and W. T. Kim, *Acta Mater*. 50, 1749 (2002).
229. Yoshizawa, K. Yamauchi, T. Yamane, and H. Sugihara, *J. Appl. Phys.* 64, 6047 (1988).
230. K. Suzuki, N. Kataoka, A. Inoue, A. Makino, and T. Masumoto, *Mater. Trans. JIM* 31, 743 (1990).
231. M. A. Willard, D. E. Laughlin, M. E. McHenry, D. Thoma, K. Sickafus, J. O. Cross, and V. G. Harris, *J. Appl. Phys.* 84, 6773 (1998).
232. K. Hono, *Progr. Mater. Sci.* 47, 621 (2002).
233. K. Hono, K. Hiraga, Q. Wang, A. Inoue, and T. Sakurai, *Acta Metall. Mater.* 40, 2137 (1992).
234. J. D. Ayers, V. G. Harris, J. A. Sprague, W. T. Elam, and H. N. Jones, *Acta Mater.* 46, 1861 (1998).
235. M. Ohnuma, K. Hono, H. Onodera, J. S. Pedersen, and S. Linderoth, *Nanostruct. Mater.* 12, 693 (1999).

236. A. Makino, K. Suzuki, A. Inoue, and T. Masumoto, *Mater. Trans. JIM* 32, 551 (1991).
237. A. Makino, A. Inoue, and T. Masumoto, *Mater. Trans. JIM* 36, 924 (1995).
238. L. K. Varga and K. V. Rao, *Nanostruct. Mater.* 12, 1157 (1999).
239. K. Suzuki, A. Makino, A. Inoue, and T. Masumoto, *J. Appl. Phys.* 74, 3316 (1993).
240. Y.-M. Moon, K.-S. Kim, S.-C. Yu, and K. V. Rao, *J. Magn. Magn. Mater.* 177, 968 (1998).
241. B. Shen, H. Kimura, and A. Inoue, *Mater. Trans.* 43, 589 (2002).
242. R. Coehoorn, D. B. de Mooij, J. P. W. B. Duchateau, and K. J. H. Buschow, *J. Phys. C* 8, 669 (1988).
243. Y. Q. Wu, D. H. Ping, B. S. Murty, H. Kanekiyo, S. Hirosawa, and K. Hono, *Scripta Mater.* 45, 355 (2001).
244. S. Hirosawa, H. Kanekiyo, and M. Uehara, *J. Appl. Phys.* 78, 6488 (1993).
245. K. Suzuki, J. M. Cadogan, M. Uehara, S. Hirosawa, and H. Kanekiyo, *Scripta Mater.* 42, 487 (2000).
246. S. Hirosawa and H. Kanekiyo, *Mater. Sci. Eng. A* 217/218, 367 (1996).
247. A. Manaf, R. A. Buckley, and H. A. Davies, *J. Magn. Magn. Mater.* 128, 302 (1993).
248. A. Inoue, A. Takeuchi, A. Makino, and T. Masumoto, *Mater. Trans. JIM* 36, 962 (1995).
249. T. Yoneyama, T. Yamamoto, and T. Hidaka, *Appl. Phys. Lett.* 67, 3197 (1995).
250. M. Sherif El-Eskandarany, K. Aoki, K. Sumiyama, and K. Suzuki, *Acta Mater.* 50, 1113 (2002).
251. A. W. Weeber, H. Bakker, and F. R. Boer, *Europhys. Lett.* 2, 445 (1986).
252. D. V. Louzguine and A. Inoue, *Mater. Res. Bull.* 34, 1991 (1999).
253. A. A. Murav'eva and O. S. Zarechnyk, *Inorganic Mater.* 6, 933 (1970).
254. D. V. Louzguine and A. Inoue, *Mater. Sci. Eng. A* 304–306, 505 (2001).
255. J. W. Christian, "The Theory of Transformations in Metals and Alloys," p. 542. Pergamon Press, Oxford, 1975.
256. D. V. Louzguine, A. Takeuchi, and A. Inoue, *J. Mater. Sci.* 35, 5537 (2000).
257. E. Yu. Ivanov, I. G. Konstanchuk, B. D. Bokhonov, and V. V. Boldyrev, *Reactiv. Solids* 7, 167 (1989).
258. F. Schurak, J. Eckert, and L. Schulz, *Nanostruct. Mater.* 12, 107 (1999).
259. V. V. Tcherdyntsev, S. D. Kaloshkin, A. I. Salimon, I. A. Tomilin, and A. M. Korsunsky, *J. Non-Cryst. Solids* 312–314, 522 (2002).
260. A. I. Salimon, A. M. Korsunsky, E. V. Shelekhov, T. A. Sviridova, S. D. Kaloshkin, V. S. Tcherdyntsev, and Yu. V. Baldokhin, *Acta Mater.* 49, 1821 (2001).
261. A. Singh and S. Ranganathan, *Mater. Sci. Eng. B* 32, 137 (1995).
262. V. S. Zhernakov, V. V. Latysh, V. V. Stolyarov, A. I. Zharikov, and R. Z. Valiev, *Scripta Mater.* 44, 1771 (2001).
263. S. Ukai and M. Fujiwara, *J. Nuclear Mater.* 307, 749 (2002).
264. H. Guckel, K. J. Skrobis, T. R. Christenson, J. Klein, S. Han, B. Choi, E. G. Lovell, and T. W. Chapman, *J. Micromechanics* 1, 135 (1991).
265. R. K. Kupka, F. Bouamrane, C. Cremers, and S. Megtert, *Appl. Surface Sci.* 164, 97 (2000).
266. P. Rai-Choudhury, "Handbook of Microlithography, Micromachining and Microfabrication 1–2." SPIE Optical Engineering Press, 1997.
267. S. Nakamura, K. Suzuki, M. Ataka, H. Fujita, H. Basrour, V. Soumann, M. Labachelerie, and W. Daniau, *Microsys. Technol.* 5, 69 (1998).
268. T. Aizawa, T. Kuji, and H. Nakano, *Mater. Trans.* 42, 1284 (2001).
269. R. B. Inturi and Z. Szklavska-Smialowska, *Corrosion* 48, 398 (1992).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanostructured Metals in Surface Enhanced Raman Spectroscopy

Helena I. S. Nogueira, José J. C. Teixeira-Dias, Tito Trindade

University of Aveiro, Aveiro, Portugal

CONTENTS

1. INTRODUCTION

Since the 19th century, a few pioneering scientists have dedicated efforts to understand the physical and chemical properties of metal colloids, in particular of gold colloids. Michael Faraday was probably the first nanochemist, in the sense that he employed chemical reasoning in the synthesis of stable gold colloids and correctly assigned their unusual optical properties to size effects occurring in particles of very small dimensions [1]. On his own words: "Agents competent to reduce gold from its solution are very numerous, and may be applied in many different ways, leaving it either in films, or in an excessively subdivided condition. Phosphorus is a very favourable agent when the latter object is in view. If a piece of this substance be placed under the surface of a moderately strong solution of chloride of gold, the reduced metal adheres to the phosphorus as a granular crystalline crust. If the solution be weak and the phosphorus clean, part of the gold is reduced in exceedingly fine particles, which becoming diffused, produce a beautiful ruby fluid" [1]. Gustav Mie is still one of the most cited authors in the scientific literature due to his work on the optics of turbid media (and metal colloids in particular) [2]. The Nobel Prize laureate in chemistry in 1925 was Richard Zsigmondy for his contributions to the foundation of modern colloid chemistry. Among numerous contributions, Zsigmondy was able to measure particle sizes of gold colloids using one of his inventions, an ultramicroscope. Other outstanding and pioneering scientific contributions could be named to demonstrate that nanometals have been a scientific issue for a long time. Nevertheless, as it is generally recognized today, the term nanometal (or nanomaterial in a broad sense) has never been so frequently used in diverse innovative contexts as it is today. There are a number of factors which may explain this upsurge on this class of materials, which include: the recognition that nanomaterials have properties different from those of the macrocrystalline analogues; the advent of characterization techniques at an atomic scale; the increase of computing power allowing the simulation of larger scale nano-objects, and the tendency to achieve better performances in miniaturized devices.

The main focus in this chapter will be on the application of nanostructured materials in surface enhanced Raman scattering (SERS) studies. The surface-enhanced Raman effect has been known since the studies of Jeanmaire and VanDuyne [3] and Albrecht and Creighton [4] on the adsorption of pyridine on a rough silver electrode. This spectroscopic method differs from other types of techniques in that the relevant physico-chemical phenomena occur at the surface of a solid phase when interacting with specific adsorbates. Moreover, it is a spectroscopic tool generally applied to very dilute samples of the adsorbate. Notably these two factors are of great interest for nanosciences and, in particular, for the chemical manipulation of metal nanoparticles and the study of its surfaces.

Nanometals are of great interest because they show unique properties related to their small size and large surface-to-volume ratio. The fabrication of nanodevices using chemical methods will be dependent on the ability to control the chemical surface in nanoparticles, which may be regarded as the building blocks for novel nanostructures. Contrary to the understanding of their bulk properties, the chemistry of nanoparticles surfaces is less understood. Therefore, a nondestructive technique, such as Raman spectroscopy, which can give information on the type of interactions occurring between molecular species

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (699–715)

and metal nanoparticles, is scientifically valuable in the context of nanotechnology.

Most of the work reported so far on the SERS effect has involved silver and gold colloids and, though in considerably less number, other noble metals such as platinum. To understand the importance of SERS in the study of metal nanoparticles and their assemblies, it is relevant to be aware of some fundamental properties of metal nanocrystals. Therefore, this chapter starts with a summary of some basic concepts involved in the synthesis, surface derivatization, and assembly methods of metal nanocrystals. These nanostructures are zero-dimensional (i.e., they correspond to systems that are finite in all directions). SERS as a Raman spectroscopic method will then be reviewed and elucidative examples will be given. Finally we will note some perspectives of development of this technique taken in consideration of its applications to nanosciences. Since the authors have a chemical background the emphasis will be on chemical nanostructures, which are nano-objects prepared by chemical methods. For more extensive discussions on nanometals the reader is referred to [5–11].

2. CHEMISTRY OF METAL NANOPARTICLES

2.1. General Concepts

Metal nanocrystals show distinct properties from those of the macrocrystalline counterparts due to their reduced particle size and larger surface to volume ratios. Since the semantics in nanosciences literature vary widely, it is important to make clear the terminology used here which follows closely the criteria adopted by Kreibig and Vollmer in defining the different classes of metal clusters [11]. These authors distinguish metal clusters according to the dimensions of the particles and they have used the general term "cluster" to define those particles. For the purposes of this chapter, the term "nanocluster" will be used here to mention nanoparticles composed of up to hundreds of atoms, with dimensions typically less than 3 nm in diameter, which includes the "very small clusters" and "small clusters" mentioned by those authors [11]. Here, the "larger clusters" will be generally referred as "nanocrystals"; this includes the larger nanoparticles forming most of the metal colloids employed in SERS which may have dimensions even superior to 100 nm. In metal nanocrystals the internal crystalline structure is similar to the macrocrystalline lattice but they show distinct physical and chemical properties from the bulk materials and from nanoclusters of metallic atoms.

Chemists have long recognized the importance of the availability of a large surface area in nanometals, for example in heterogeneous catalysis [12]. Surface chemistry controls to a large extent the properties of catalysts, quantum semiconductor devices, and bioprobes, to mention just a few examples. Due to the large surface energy, nanoparticles are thermodynamically unstable and with time tend to coalesce into the stable bulk structure. To prevent agglomeration, the metal surfaces are molecularly tailored using Lewis bases which render to the nanoparticles steric hindrance. Unless otherwise stated, it is assumed that the metal nanocrystals and nanoclusters reported here are generated by chemical means leading to surface protected nanostructures. From a chemical perspective there is an obvious resemblance between the nanoclusters and the larger nanocrystals, in the sense that organic ligands are used to stabilize the metal core.

It is now well established that the intrinsic properties of metals change as the particle size is reduced from the bulk material to the nanocluster scale. A few selected examples will be given which clearly elucidate this important point. A fundamental aspect, which immediately arises, is related to the structure development from an assembly of a few metallic atoms in a nanocluster to a typical macrocrystalline lattice. Although this topic is still a matter of intense investigation there are some related important points to be addressed. First, it should be kept in mind that nanoclusters are in fact quantum-sized objects. Therefore their stable structures are governed by quantum mechanics; *ab initio* methods have been employed to predict the structures of small nanoclusters [11]. The structural patterns observed in very small nanoclusters may no longer be related to the typical geometrical arrangement of the atoms observed in the bulk lattice. Schmid and co-workers have reported a series of giant metal nanoclusters whose cores correspond to geometrical arrangements containing a specific number of atoms, often called magic atom-number clusters [6, 7]. These nanoclusters seem to be built by the progressively addition of shells of identical atoms to a central atom. For example, it has been shown that gold clusters, which in the bulk crystallizes with a face centered cubic lattice, adopt specific geometrical arrangements depending on the number of atoms forming the nanocluster [6, 7]. The experimental identification of metal nanoclusters with a specific number of atoms (magic atom-numbers) has also been reported for example for sodium and gold [13, 14].

The structural transformations occurring as the particle size reduction proceeds down to the nanocluster regimen are accompanied by dramatic changes in the electronic, magnetic, catalytic, and thermodynamic properties [10]. A first effect of reducing particle size is the creation of more surface sites. An elucidative example of these changes is exemplified by the dramatic decrease on the melting temperature of metals as the particle size is reduced to a few nanometers. The melting temperature of metal nanoclusters decreases with particle size and the rate of decrease increases drammatically at very small sizes [10]. In terms of the electronic structure of the metal nanoclusters, the energy levels forming a continuum half-filled conduction band in a bulk metal become more discrete, approaching the molecular-like description based on more localized molecular bonds [8, 11]. In these nanoclusters very strong size effects govern their electronic and structural properties; therefore theoretical models derived from quantum mechanics should apply to have a complete understanding of this behavior. Measurements on the ionization potentials show that depending on the number of metal atoms composing the nanoclusters there are specific sizes and geometries which are the most stable; as mentioned these correspond to specific atom-numbers. Due to confinement of the electron wavefunction by the particles dimensions, the bands in these systems become best described as sets of discrete energy levels [8].

The average spacing within successive quantum levels is known as the Kubo gap and is given by

$$\delta = 4E_F/3n \quad (1)$$

where E_F is the Fermi energy of the bulk metal and n is the number of valence electrons in the nanoclusters [8]. The properties of nanoclusters, such as electrical conductivity and magnetic susceptibility, may present quantum size effects due to the existence of the Kubo gap. In certain conditions, low temperatures for instance, the metal nanoclusters may exhibit nonmetallic behavior since this gap is comparable to the thermal energy necessary to promote electrons across successive energy levels.

SERS studies on nanostructured metals are usually performed on metal colloids whose dimensions are typically in the 10–200 nm range. These nano-objects have dimensions larger than those of the nanoclusters mentioned previously and show extrinsic effects, which are size-dependent responses to external fields or forces irrespective of the intrinsic effects observed for nanoclusters [11]. The optical properties of metal nanocrystals and nanoclusters are size-dependent and distinct from the bulk metals. For example, it is well known that macrocrystalline metals are highly absorbent in the infrared region of the electromagnetic spectrum; in contrast metal nanoparticles are transparent in this wavelength region. Gold and silver nanoparticles, the most used metals in SERS studies, have characteristic optical spectra [16, 17], showing strong bands in the visible (Fig. 1). The incidence of ultraviolet (UV)/visible light in metal nanocrystals originates an absorption band due to the spatial confinement of the electrons of the conduction band. Due to this confinement effect, the plasmon oscillations become restricted to a narrow frequency range located in the visible [11]. The optical spectra of metal nanocrystals can be predicted using the Mie equations. For isolated spherical particles with dimensions much smaller than the wavelength of the incident light, the absorption band due to the dipolar surface plasmon is observed at a frequency ω_0 such that [2, 9, 11]

$$\varepsilon_1(\omega_0) = -2\varepsilon_m \quad (2)$$

where

$$\varepsilon(\omega) = \varepsilon_1(\omega) + i\varepsilon_2(\omega) \quad (3)$$

$\varepsilon(\omega)$ is the dielectric function of the metal particles and ε_m is the dielectric function of the surrounding medium. By varying the nanoparticle dimensions there are shifts on the resonance band compared to the classical Mie frequency which can be interpreted using Mie's theory [11]. From a chemical point of view, it is of interest to consider surface coated metal nanocrystals. The optical spectra of coated nanoparticles can be obtained using the Mie formalism and the appropriate methodologies [9]. The ligands at the nanoparticles surface form a dielectric cap whose electric insulation can be varied for instance by varying the length of the alkyl chains of the ligand.

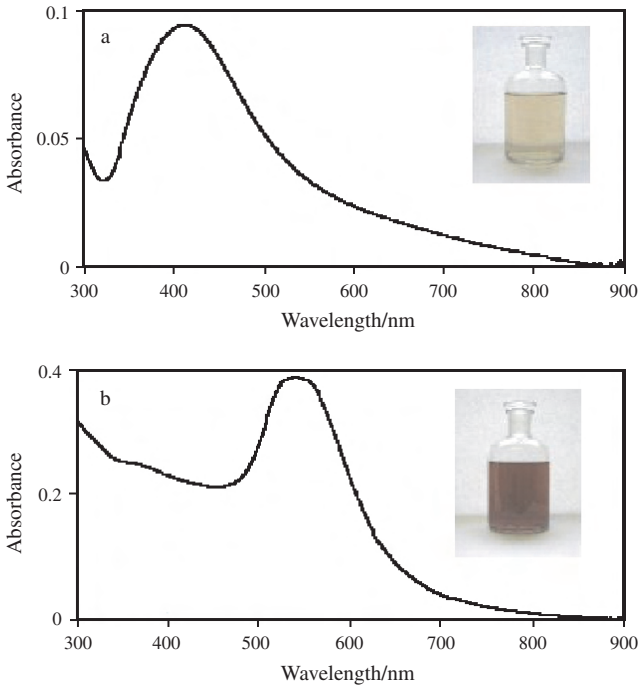


Figure 1. Absorption by silver (a) and gold (b) hydrosols.

2.2. Chemical Synthesis

Although the synthesis of metal colloids has been investigated since the work of Faraday, the interest in chemical methods to yield reasonable amounts of functional and highly stable nanometals is recent. Within the variety of synthetic methods developed during the last two decades, the need to prepare nanoparticles which are morphologically uniform, have the ability for surface functionalization and may be further used to fabricate nanostructured devices becomes clear. In general, metal nanoparticles have been prepared using the chemical strategy delineated by pioneering colloidal chemists (i.e., the addition of a reducing agent to a solution containing the metal salt). These methods vary essentially on the type of the reducing agent used, the solvent medium, the order/amount of reagent addition, and the chemical conditions employed to stabilize the as-prepared nanoparticles.

The preparation of gold nanoparticles using the citrate method, still one of the most popular chemical methods to obtain gold colloids, involves the reduction of the $[AuCl_4]^-$ complex by citric acid [18]. Similarly, Ag colloids can be prepared by this method using silver nitrate or acetate as the starting salt. The metal particles are electrostatic stabilized in solution, during a relatively long period of time, due to a citrate rich surface which avoids nanoparticle coalescence. In aqueous solution, electrostatic stabilization of metal nanocrystals has been commonly used. The stability and quality of these colloids for SERS can be easily evaluated by their optical spectra, which should be dominated by the characteristic surface plasmon sharp bands (Fig. 1). Broadening of the bands is associated with agglomeration processes. In some situations this agglomeration step might be important to record good SERS spectra. The SERS activity/sensitivity of these nanoparticles is strongly dependent on their morphological characteristics. In other words, in

a nanoparticle ensemble, not all the particulates will have the same SERS sensitivity. This was shown by investigating the SERS activity of silver colloids which had been size fractionated from a morphologically heterogeneous sample [19]. In a typical silver colloid prepared by the standard citrate-reduction procedure of Lee and Meisel, only about 0.1–1% of particles are highly SERS sensitive. These particles are large faceted nanocrystals (80–100 nm) which can be concentrated in a colloid to about 10–15% in relation to the starting sample, using size-selective fractionation methods [19]. Morphological modifications on the generally spheroidal metal nanoparticles may be obtained by modifying the chemical speciation during the synthesis. Gold nanorods have been reported via photochemical synthesis which involved reducing $[AuCl_4]^-$ in the presence of $AgNO_3$ and organic surfactants [20]. The gold nanorods have diameters of 12.6, 15, and 18 nm, depending on the amount of silver nitrate used during the synthesis. Gold nanoparticles spherical in shape were mainly obtained in the absence of silver ions [20].

The synthesis of metal nanocrystals can be carried out in a solvent medium which acts as the reducing agent of the metal salt [21–23]. Organic solvents such as ethanol [21], ethylene glycol [22], and dimethylformamide [23] have been used to obtain metal nanocrystals with distinct particle sizes and morphological properties. These methods involve some sort of external energy (e.g., heat) to promote the reduction of the metal salt. It should be noted that although SERS studies on metal colloids involve frequently an aqueous medium, there are reports on SERS spectra obtained in other solvents [24]. Good quality SERS spectra of organic adsorbates have been reported, using silver colloids dispersed in organic solvents such as dimethylformamide and acetonitrile [24].

In the chemical synthesis of metal nanocrystals, experimental conditions are set up so that the particle growth process is terminated abruptly and a stable sol is formed in the liquid medium. The stability of the as-formed particles against agglomeration can be attained by surface capping or by electrostatic repulsion. This general idea is behind many synthetic methods in which polymers, surfactants, or Lewis bases have been used to surface cap the small particles formed (nanocrystals). An elucidative example of this concept is the reduction of a metal salt in confined spaces such as reverse micelles [25]. Reverse micelles have nonpolar tail groups forming an interface with the bulk hydrocarbon solvent. The polar head groups are directed toward the interior of the micelle, which contains a droplet of a polar solvent, usually water. The formation of the metal phase occurs inside this nanoreactor by reduction of a metal salt. The aqueous droplet is capable of solubilizing the starting metal salt and also provides the medium in which nucleation and growth occurs during the reduction of the metal salt. In the context of this chapter, it is important that the metal nanocrystals surfaces can be modified using either organic ligands or a second inorganic phase to replace the surfactant molecules. One of the major advantages of metal reduction inside reverse micelles is its simplicity associated to two important requirements on metal nanoparticle preparation. These are (i) size-tuning and size polydispersity control (by varying $w = [\text{water}]/[\text{surfactant}]$) and, (ii) surface modification, since the surfactant molecules at the nanocrystals surface can be replaced by other chemical capping agents, such as thiols or amines. A simple relation was deduced to predict the water pool radius in reverse micelles using bis(2-ethylhexyl)sulfosuccinate (AOT) as the surfactant [25]:

$$R_w(\text{Å}) = 1.5w \tag{4}$$

The use of confined spaces in organized amphiphilic structures has been employed in the synthesis of several metals in quite distinct contexts. These include the formation of ultrathin particulate Ag films [26], local control of the shape of Cu particles [27, 28], and the formation of monodispersed nanoparticles for a wide range of metals, including bimetallic systems [25, 29]. In these few examples the subject of study is related to the metal nature surface and in particular to the type of chemical bonding occurring between the capping agent and the metal sites. It should be noted that usually the capped nanometals can be isolated as powders and redispersed in solvents for further chemical or physical treatments. In order to obtain narrow particles size distributions the nanodispersed metal can be further subjected to size-selected precipitation methods [30].

Other confined spaces have been used in the preparation of metal nanoparticles. Polymers, for example, are interesting stabilizers to produce metal nanoparticles with controlled shape and size. The reduction of $[AuCl_4]^-$ to yield gold nanoparticles has been performed at room temperature in the presence of poly(vinylpyrrolidone) as the stabilizer using the oxygen-free solvent formamide [31]. As mentioned, this is a situation in which the solvent formamide is the reducing agent. Nonspherical Pd nanocrystals have been prepared using surfactant micelles and a polymer gel [32]. In this regard, the UV photolysis of $Pd(acac)_2$ previously solubilized in anionic SDS (sodium dodecyl sulfate) micelles and poly(acrylamide) gel yielded faceted Pd nanocrystals (tetrahedral or octahedral shaped). Although the mechanism of formation of these Pd nanoparticles is not known, there is a synergistic effect involving both the micelles and the polymer, in a process possibly mediated by the type of interaction occurring at specific metal surfaces with the chemical sites of the polymer matrices. The polymer gel confers some rigidity to the confined spaces limiting the shape changes occurring on the SDS micelles due to thermal motion. The synthesis in the absence of the polymer did not yield nonspherical Pd nanocrystals, revealing the importance of the polymer gel to influence the faceted final morphology of the particles. Fine divided Pd is a very important heterogeneous catalyst; the knowledge of which surfaces are active in chemical catalysis may be inferred during the synthetic process, including *in-situ* spectroscopic analysis of the particle formation. Therefore, Raman spectroscopy is a potential tool to investigate the interaction of well-defined Pd nanoparticles, as well as other faceted noble metal particles, with adsorbates in other media, including hard matrices such as pillared clays [33, 34], metal oxides [35], or silica spheres [36]. Synthesis in large quantities is also a major limitation on most of the chemical methods employed to produce nanometals. Although for SERS applications scale-up synthesis is not needed, this point should

be considered for the use of nanometals in the large-scale production of devices or in catalytic processes. A possible strategy to overcome this limitation involves the synthesis of nanometals in confined spaces of solid polymer matrices. The matrix acts as an active template preventing nanoparticle agglomeration and promoting surface passivation. An important aspect in this strategy is the possibility to redisperse the nanometals allowing further chemical manipulation. This method has been employed in the preparation of Pt nanoparticles in the solid state via a polymer matrix [poly(*N*-vinyl-2-pyrrolidone)] effect [37].

The reduction of metal salts in order to produce nanosized metals has also been accomplished using other strategies and a few more will be briefly mentioned. Copper nanoparticles have been prepared by thermal reduction and sonochemical reduction of the precursor copper(II) hydrazine carboxylate [38]. Despite the use of the same copper precursor, under an argon atmosphere, the chemical composition of the final products differs depending on the method used. Hence, whereas the thermal reduction yielded pure irregularly shaped Cu nanoparticles (200–250 nm), a mixture of the metal phase and Cu_2O was obtained when the sonochemical method was applied. The presence of the oxide phase was attributed to the generation of OH· radicals that are formed during the sonolysis of water. The sonochemical synthesis under a mixture of argon and hydrogen atmosphere led to pure aggregated copper nanoparticles (50–70 nm). This work also highlights the importance of the preparative method for the final properties of the nanometals, and their use in processes in which the surface plays an important role, such as catalysis. In fact, the thermal derived nanoparticles are faceted and dense while the sonochemical derived nanoparticles are porous. Both syntheses yielded Cu nanoparticles showing superior catalytic activity (Ullmann reaction) as compared to the analogous commercial powders. Copper and palladium form alloys over their entire composition range. Bimetallic palladium–copper nanoparticles stabilized by poly(vinylpyrrolidone) (PVP) have been prepared by thermoreduction (135 °C) of Pd and Cu acetates in 2-ethoxyethanol [39]. Poly(vinylpirrolidone) was used here to surface cap the metal nanoparticles allowing the formation of highly monodispersed PdCu/PVP bimetallic particles (30–50 Å ± 10%).

Electrochemical methods have been shown to meet two important criteria in nanoparticle synthesis, particle size control and surface stabilization [40]. Hence, tetraalkylammonium-stabilized Pd and Ni nanoparticles have been reported by a simple electrochemical method. Size tuning is achieved by variation of the current density and high yields of the nanometals are obtained [40]. The formation of the metal powders is avoided due to the presence of tetraalkylammonium or phosphonium salts. The observed particle sizes are in accordance with the equation for the critical radius (R_{crit}) of adatoms formed at the cathode interface [40],

$$R_{crit} = \frac{2M\gamma}{nF\eta\rho} \tag{5}$$

where M is the molecular weight, γ is the surface tension, F is the Faraday constant, η is the overpotential, ρ is the density of the nanoparticles, and n is the valency. Equation (5) shows that the particle critical radius is inversely dependent on the applied overpotential and therefore the current density. Note that numerous studies on SERS have been performed on metal electrodes and this type of chemical synthesis may constitute an interesting bridge to nanoparticle formation and mechanisms for surface detection.

Rigorous size tuning and surface tailoring are still the major challenges in the chemical synthesis of metal nanoparticles. The delineation of diverse chemical strategies to provide useful alternatives in the synthesis of various metals is still a very active research area. Associated with the chemical synthesis of nanometals, there are in some cases postpreparative methods, such as size-fractionation procedures and surface modification methods. The main idea is to end up with a quasi-monodispersed nanometal with functionalized surfaces. These particles will form the basis of devices in which the properties related to the dimensions and high surface area of the building units are fully exploited. The use of nanocomposite particles containing a SERS active metal as one of the components should also be pointed out. As will be seen later this may prove interesting to expand the studies of SERS to various types of surfaces.

2.3. Surface Modification and Assembly

The ability to control the chemical surface in nanometals is essential to the fabrication of nanodevices. This is directly related to the use of assembly methods to produce organized nanostructures using highly stable passivated nanocrystals, here regarded as the building blocks. The surface derivatization of the metal particle surface should not disturb the integrity of the metal core. The particle surface can be derivatized by either an inorganic layer, an organic layer, or both. Only the organic surface derivatization of nanometals will be mentioned here. The silica encapsulation of semiconductor and metal nanocrystals has been reviewed [41]. The type of organically capped surface allows the control of the geometrical arrangement and spacing of the nanometal particles, as well as of the higher order structure observed in the final assembly.

Thiol or ammine passivated metal nanocrystals have been the most investigated organically capped systems. The surface passivation of gold nanoparticles with thiol compounds (e.g., alkanethiols with long alkyl chains) has been widely used since the work of Brust and co-workers [42, 43]. These authors have reported the formation of dodecanethiol passivated gold nanoparticles by reducing $[AuCl_4]^-$ in a toluene solution, using aqueous $NaBH_4$ as the reducing agent, after phase transfer using the cationic surfactant tetraoctylammonium [42]. The high affinity of the alkanethiol molecules to the gold surfaces prevents the coalescence of the gold nanoparticles into the bulk phase and also allows the isolation of the nanometal as a powder, which is redispersible in a wide range of organic solvents. The overall reaction is summarized by [42]

$$AuCl^-_{4\,(aq)} + N(C_8H_{17})^+_4(C_6H_5Me) \longrightarrow N(C_8H_{17})^+_4AuCl^-_4(C_6H_5Me) \tag{6}$$

$$m\mathrm{AuCl_4^-(C_6H_5Me)} + n\mathrm{C_{12}H_{25}SH(C_6H_5Me)} + 3me^- \longrightarrow 4m\mathrm{Cl^-_{(aq)}} + (\mathrm{Au}_m)(\mathrm{C_{12}H_{25}SH})_n(\mathrm{C_6H_5Me}) \quad (7)$$

By varying the length of the alkyl chain, the diameter of the gold nanocrystals can be varied within the nanosize regimen. This type of Au nanocluster contains 200–1000 atoms as determined by mass spectrometry and electron microscopy [14, 44]. The aurothiol nanoclusters reversibly condense to form films or powders [14, 44]. An interesting feature of these Au nanoclusters relies on the possibility of self-assembly into ordered close-packed structures due to steric hindrance of the long alkyl chains. X-ray diffraction performed on these superstructures revealed that in addition to the diffuse peaks due to the Au atoms within individual nanoclusters, there are sharp features in the small angle region attributed to three-dimensional (3D) superlattice formation [44]. The nanoclusters assemblies seem to consist of compact crystalline Au cores capped at the surface by thiolates bound to Au sites through the sulfur atoms. The *n*-alkyl chains tend to organize and fill the voids within the nanoclusters to obtain maximal dense packing. The observed high symmetry structures, hexagonal close-packed and cubic close packed, observed for the 3D superlattices become destabilized as the *n*-alkanethiol monolayers become thicker in relation to the core dimensions, leading to more open structures such as the body-centered cubic structure. This has been described as an increase of the softness of the Au nanoparticles which results in superstructures of lower symmetry and lower coordination numbers [44]. These nanostructures are of great interest not only from the academic point of view but also to fabricate nanodevices, since they can be functionalized and integrated into other surfaces and solid state structures [45, 46]. There have been various adaptations to the method initially proposed by Brust et al. to produce thiol-cappped nanometals. One interesting example consists of the formation of nanoparticles and the surface passivation with thiols in a one-pot method [47]. Such a method consists on the preparation of a hydrosol of Au (it works also for Ag or Pt) nanoparticles, by reducing $[AuCl_4]^-$ with tetrakis(hydroxymethyl) phosphonium chloride. A toluene solution containing the alkanethiol is added to the hydrosol; further addition of HCl to the biphasic mixture led to the migration of the metal nanoparticles to the organic phase. This is also an interesting experiment to show the aurifilicity of sulfur compounds.

The self-assembly of *n*-alkanethiolate monolayers has been used on other metal nanostructures, for example silver nanocrystals [48, 49]. These monolayers encapsulate the Ag nanoparticles, prepared electrochemically, protecting the metal sites from further reactions for example by the anodic dissolution in pure water. Using this strategy, self-assembly and chemical protection have not been observed in using long-chain *n*-alkyl amines [49]. However, these compounds have been used as efficient surface passivating agents on gold nanoparticles (25–70 Å) [50]. These amine-capped Au nanocrystals have been prepared by the reduction of $[AuCl_4]^-$ by $NaBH_4$ in the presence of a series of *n*-alkylamines. Therefore, a method similar to the original method proposed by Brust et al. was employed with the main exception that *n*-alkylamines were used instead of *n*-alkanethiols. The amine/gold surface interaction is charge-neutral via weak covalent bonds. The hydrophobic character of the amine-capped Au nanoparticles allows dispersion in a range of organic solvents although this behavior strongly depends on the synthetic method used in their preparation [50].

In principle, hybrid superlattices composed of metal and semiconductor nanocrystals could be obtained by the assembly methods employed for the single components. The appropriate tailoring of the nanocrystal surfaces should be viewed as an influential parameter for obtaining well-organized hybrid nanostructures. For example, layered nanocomposites of close-packed gold nanoparticles and TiO_2 layers have been reported [51]. The gold nanoparticles were prepared by reduction of $HAuCl_4$ by $NaBH_4$ in the presence of the disulfide $HO(CH_2)_{11}SS(CH_2)_{11}OH$, which acts as the nanocrystal surface stabilizer. The modified gold nanocrystals have surface hydroxyl groups being chemically compatible to participate in the sol–gel process used to generate the TiO_2 particles from metal alkoxide hydrolysis. A general procedure to obtain hybrid superlattices, of metal–metal or metal–semiconductor type, has been reported [52]. These heterostructures may consist of alternate layers of semiconductor (CdS) and metal (Pt) nanocrystals obtained from layer-by-layer deposition by using dithiols as cross-linkers.

The mechanism of assembly of nanoparticles involves nanocrystal-specific interactions still not fully elucidated, both in terms of capping agent/nanoparticle interactions and interparticle (surface capped) interactions. Hence, the amine-capped Au particle stability with respect to agglomeration was attributed to kinetic factors [50] whereas the alkane-capped Au particles were shown to possess thermodynamic stability. With respect to the interparticle interactions involved in the formation of larger nanostructures, besides the type of capping agent, the possible mechanism seems to depend on the exact nature of the particle core. In the formation of thiol capped Ag nanocrystal superlattices [53], the capped inorganic nanocrystals were described as soft balls, the organic ligands filling the voids between the Ag cores, hence determining the interparticle spacing. Entropy effects at high particle volume fractions drive the well positioning of the capped Ag nanocrystals into an extended face centered cubic structure. More recently, it was reported that the incorporation of carboxylic groups in thiolate capping agents allows controllable interparticle packing, ordering, and branching of Au nanostructures [54]. A model involving both hydrogen bonding at the carboxylic groups and cohesive van der Waals interdigitation of the thiolate molecules was proposed to explain the interparticle assembling forces.

A number of methods have emerged recently in order to produce metal nanoparticles with a variety of functional groups other than thiolates. Although the fundamental goals on these procedures are generally similar, [i.e., surface tailoring for nanometal protection and further assembly into larger (and distinct) nanostructures], the fact is that depending on the type of capping, chemical selectivity can be attained. A fundamental idea in this approach aims to make use of the particular molecular shape of the cap to bind to metal sites. Calix[4]arene type molecules have been particularly investigated as metal receptors [55, 56]. Gold-thiol

monolayers of calix[4]resorcinarenes for molecular recognition and sensing have been used. For example the calix-4-resorcinarene molecules can anchor to the gold surface using the thiol terminated alkyl chains, leading to well-packed monolayers (each 22–24 ± 2 Å) at 60 °C [55]. In these monolayers the bowl type structure formed by the phenol rings may act as molecular receptors for specific adsorbates. Other examples of the preparation of nanostructured functional materials due to chemical selectivity are the preparation of cyclophane receptor–Au nanoparticles for electrochemical sensing [57] or the photoinduced charge separation in fluorophore–gold nanoparticles [58]. Table 1 summarizes some systems in which organic compounds have been used to surface passivate metal nanoparticles in order to form functional superstructures. It should be noted that SERS studies have rarely been used in the characterization of these hybrid systems which open up interesting possibilities to develop this spectroscopic method in the field of the nanosciences.

3. SURFACE ENHANCED RAMAN SPECTROSCOPY (SERS)

3.1. Fundamentals of Raman Spectroscopy

In the issue of *Nature*, March 31, 1928, a scientific note entitled "A new type of secondary radiation" was published, signed by Raman and his student Krishnan [60]. Using an analogy with X-ray scattering, it was claimed that "also in the case of ordinary light two types of scattering, one determined by the normal optical properties of the atoms or molecules, and another representing the effect of their fluctuations from their normal state" should be expected. A simple experiment with a beam of sunlight, a liquid sample in between two light filters—a blue–violet filter and a green–yellow one—and a light detector supported those conclusions. In a subsequent issue of *Nature* (April 21), another paper entitled "A change of wavelength in light scattering" was published, this time signed only by Raman [61]. It is interesting to note the worries of Raman at the time and his reasoning to convince himself that the observed secondary radiation was not a fluorescence. The "excessive feebleness" of this "modified scattering" and its "polarization characteristics" were pointed out. A month later, in yet another issue of *Nature* (May 5), Raman and Krishnan published an explanatory note of the observed effect, entitled "The optical analogue of the Compton effect" [62].

Table 1. Examples of organic ligands used in nanometal passivation.

Compound	Nanometal (typical dimensions)	Ref.
Dodecylamine ($C_{12}H_{25}NH_2$) Oleylamine ($C_{18}H_{35}NH_2$)	Au (2.5–7 nm)	[50]
Calyx[4]resorcinarenethiols	Au (film)	[55]
	Au (film)	[56]
Dodecanethiol	Ag (<10 nm)	[53]
Decanethiol+α, ω-mercapto-alkanoic acids	Au (6.2 ± 0.5 nm)	[54]
Pyrene+dodecathiol	Au (2–3 nm)	[58]
C-undecylcalix[4]resorcinarene	Co (27 ± 4 nm)	[59]
Decanedithiol	Au (6 nm) and Pt (5 nm)	[52]

When monochromatic light of frequency ν_0 is incident on a transparent sample and the scattered light is dispersed by a monochromator and subsequently measured in intensity, the spectrum of inelastically scattered radiation, the Raman spectrum of the sample, is obtained. A dispersive instrument for Raman spectroscopy includes a source of monochromatic radiation, a sample device, a dispersing system, and a detection/recording system. An adequate choice of the frequency of the laser is required to fulfill the objectives of the experiment. In particular, fluorescence may be avoided if a sufficiently long laser wavelength, preferably 752.5 nm, is selected [63]. Modern Raman spectrometers are usually equipped with both a macro sample compartment and a microscope. In fact, so important is the possibility of studying microscopic samples that Raman microspectroscopy is a new experimental field on its own [64]. The dispersing system consists of a monochromator to disperse the light scattered by the sample. Through scattering at optical surfaces, particularly at the gratings, a small fraction of light at the laser frequency appears in the output focal plane, at positions corresponding to other frequencies in the neighborhood of the excitation frequency. In order to improve stray light rejection, double or preferably triple dispersion should be used [65]. At the detection stage, monochannel and multichannel detectors can be employed. Comparison of the quantum efficiency curves of a photomultiplier tube and a charge-coupled device, in the studied wavelength region, should be done before one of these detectors is selected [66].

The total scattering intensity I_S in photons per molecule per second, for a transition from the initial ground state, $|G\rangle$, to the final state, $|F\rangle$, is given by

$$I_S \propto \omega_S^4 I_L \phi(\omega_S, \Gamma) \Sigma |(\alpha_{\rho\sigma})_{GF}|^2 \tag{8}$$

where the symbol $\propto$ should be read "proportional to," indicating that uninteresting constants were intentionally omitted; I_L is the incident intensity at the laser frequency ω_L, ω_S is the scattered frequency, $\phi(\omega_S, \Gamma)$ is the bandshape function, and $(\alpha_{\rho\sigma})_{GF}$ is the polarizability tensor for the transition from $|G\rangle$ to $|F\rangle$, with scattered and incident polarizations indicated by ρ and σ, respectively, and the summation extending over the polarization subscripts. Use of perturbation theory shows that the second order perturbation expression for $(\alpha_{\rho\sigma})_{GF}$ is approximately given by

$$(\alpha_{\rho\sigma})_{GF} \propto \Sigma \langle F|\rho|I\rangle\langle I|\sigma|G\rangle(\omega_{GI} - \omega_L - i\Gamma)^{-1} \tag{9}$$

where Γ is the linewidth of the intermediate state $|I\rangle$ and the summation extending over intermediate states. In (2), the term associated with the $|I\rangle$-to-$|F\rangle$ downward transition—the nonresonance term—was neglected, because its denominator, $(\omega_{IF} + \omega_L)^{-1}$, which excludes the possibility of resonance due to the plus sign, is always comparatively large.

For convenience, the scattering may be classified into three excitation frequency regions: normal (off-resonance) *Raman scattering*, for incident frequencies well away from resonance with any allowed transition; *discrete resonance Raman scattering* (DRRS), for excitation near or in resonance with a discrete transition; *continuum resonance*

Raman scattering (CRRS), for excitation resonant with continuum transitions, that is, excitation above a dissociation limit or into a repulsive electronic state [67]. In CRRS, the summation in (9) should be replaced by an integration, and the corresponding density of states of the continuum, $\rho(\omega_i)$, used.

After substitution of (9), expression (8) enables us to assess the effect of varying the frequency of the exciting radiation, thus accounting for the scattering properties of the sample. In particular, when the laser frequency approaches resonance with a real discrete transition, a single excited rovibronic state is discriminated against all the others, since the denominator in (9) is drastically reduced for that particular rovibronic state as the intermediate state and the polarizability value increases abruptly. In DRRS, the overtones may be of intensity comparable to the fundamental transition and, in fact, as strong as the Rayleigh line. The ratio of the intensity of one overtone to the next varies erratically and a small shift in frequency of the excitation line may cause drastic intensity variations. On the other hand, in CRRS, a small shift in frequency will cause a gradual variation in intensity as, within a continuum, a small frequency shift will not destroy the resonance condition.

When the intensity of a given Raman mode is plotted against the excitation frequency, an excitation profile is obtained. If this profile is compared with the absorption spectrum of the molecule, the agreement is usually good. However, in some cases, the peak of the excitation profile exhibits a redshift from the absorption maximum, as for the fundamental vibration mode of Br_2 isolated in a rare-gas matrix [68]. This effect has been studied and explained through consideration of the excitation frequency dependence of the factors involved in the intensity of a Raman transition as compared with the intensity of an absorption transition [68].

The bandshape function in (1) can be represented by the general expression

$$\phi(\omega_S, \Gamma) = \delta(\omega_L - \omega_S - \omega_{GF}) + L(\omega_S, \Gamma) \qquad (10)$$

where the L-function represents a Lorentzian, the linewidth is given by

$$\Gamma = \gamma_n + \gamma_e + \gamma_i \qquad (11)$$

and subscripts n, e, and i refer, respectively, to natural linewidth and elastic and inelastic collisions [67]. At sufficiently low pressures, both γ_e and γ_i are negligible, and the second term in (10) almost vanishes. On the other hand, if the pressure increases, the second term in (10) becomes progressively important and a Lorentzian bandshape is then obtained. Hence, a pressure increase in DRRS leads to a Γ increase and, at sufficiently high pressures, eventually leads to a change from DRRS to CRRS, where the denominator in (9) no longer discriminates against nearby states [69, 70]. It is worth pointing out that when a phase interruption occurs in the excited state, the underlying unit of the Raman process no longer prevails and one may then speak of an absorption followed by an emission (i.e., fluorescence) [67].

When the frequency of a single-mode 514.5 nm incident laser was tuned away from resonance with a discrete transition in molecular iodine, a significantly shortened scattering lifetime was observed [71]. This experiment used incident laser pulses of rectangular shape with time widths of ≈100 ns. The observation of a shortened scattering time is an indication of what is conventionally called a normal Raman process. In this case, an upper limit to the re-emission time can be estimated from

$$(\Delta\omega^2 + \Gamma^2)^{-1/2} \qquad (12)$$

For a 647.1 nm laser wavelength ($\omega_L \approx 2911$ Trad s^{-1}) and a sample with gaseous molecular iodine ($\omega_{GI} \approx 3315$ Trad s^{-1}, i.e., 568.2 nm on-resonance wavelength), $\Delta\omega \approx 404$ Trad s^{-1} and $1/\Delta\omega \approx 2.5$ fs [Γ in (12) is negligible]. On the other hand, if the laser frequency is tuned with a discrete transition in molecular iodine, a re-emission lifetime of the order of 10^3 ns is obtained [71].

3.2. The SERS Effect

Surface-enhanced Raman scattering refers to the observation that for certain molecules adsorbed on specially prepared metal surfaces a Raman spectrum is obtained whose intensity is enhanced by a factor of up to 10^6 [72]. For most surfaces the intensity of Raman bands from the surface monolayer is very low and is normally undetectable. If the surface is a suitably roughened (e.g., roughened electrodes) or there is a finely divided surface (e.g., colloids) of copper, silver or gold, however, the SERS effect then amplifies the Raman scattering, giving easily detectable Raman signals specifically from the first monolayer. The SERS effect was discovered in 1977 [3, 4] and there was strong research interest in the 1980s [73], which settled until another more recent burst of interest came out in association with research on metal nanoparticles and application to single-molecule detection [74, 75].

The large enhancement factors in SERS are obtained through a combination of an electromagnetic effect and chemical interactions between the adsorbate molecule and the surface [76]. The nature of the signal enhancement has been thoroughly studied but aspects such as the relative contribution of the electromagnetic and chemical enhancements are not fully understood. The electromagnetic effect depends on the response of the conduction electrons of the metal surface to the electromagnetic field of the incident and scattered light, being related to the excitation of localized surface plasmons [76–78]. The chemical enhancement mechanisms are related to the formation of a molecule–surface complex and charge transfer transitions between the surface and the adsorbed molecule [76, 79]. It is this chemical effect that gives the high degree of specificity of SERS to the first monolayer.

The scattering intensity in normal Raman scattering, I_{NRS}, is proportional [75] to the Raman cross section, σ^R_{free}, the excitation laser intensity, $I(\omega_L)$, and the number of molecules in the probed volume, N,

$$I_{\mathrm{NRS}}(\omega_S) \propto N \cdot I(\omega_L) \cdot \sigma^R_{\mathrm{free}} \qquad (13)$$

where the symbol $\propto$ should be read "proportional to." The surface-enhanced Raman signal, I_{SERS}, is proportional [75] to the Raman cross section of the adsorbed molecule, σ^R_{ads}, the excitation laser intensity, $I(\omega_L)$, and the number of molecules which are involved in the SERS process, N':

$$I_{SERS}(\omega_S) \propto N' \cdot I(\omega_L) \cdot |A(\omega_L)|^2 \cdot |A(\omega_S)|^2 \cdot \sigma^R_{ads} \quad (14)$$

The field enhancement factors at the laser and Stokes frequency are given by $A(\omega_L)$ and $A(\omega_S)$, respectively. The chemical enhancement is included in σ^R_{ads} which describes the increased Raman cross section of the adsorbed molecule. The SERS effect allows one to overcome the extremely small cross section for Raman scattering, typically 10^{-31}–10^{-29} cm^2/molecule (the large values occurring only under resonance Raman conditions). Raman enhancement factors on the order of 10^{14} made possible the optical detection and spectroscopic studies of single molecules and single nanoparticles, at room temperature with the use of SERS [80, 81].

3.2.1. Surface Selection Rules

The relative intensities of different vibrational modes in the Raman spectrum of a molecule will be dramatically changed when it is placed near a metal surface; the intensity of some modes may even be almost totally suppressed. This phenomenon is usually referred to as a "surface selection rule." The orientation of molecules adsorbed at metal surfaces can be deduced from the observed intensity distribution of SERS bands, based on surface selection rules [82, 83]. A number of surface selection rules for SERS have been proposed, in particular by Creighton [78, 84] and Moskovits et al. [76, 85–87]. The least enhanced vibrational modes are those whose Raman tensor components involve the two axes in the plane of the surface [76, 78]. The relative enhancements for normal modes can be classified into three groups in terms of the polarizability tensor elements. Normal modes associated with the tensor element α_{33} are the most enhanced, modes associated with α_{13} or α_{23} are less enhanced, and modes associated with α_{11}, α_{12}, or α_{22} are the least enhanced. The subscripts 1 to 3 refer to the orthogonal axes parallel (1 and 2) and perpendicular (3) to the surface.

A set of generally accepted rules is usually followed by authors to judge the molecular orientation on the surface from SERS. The rules presented next are used in the particular case of substituted benzenes.

For substituted benzenes, the presence or absence of the benzene ring C–H stretching vibration has been used [83, 88, 89] as an indicator in SERS for the perpendicular or parallel orientation of the benzene ring with respect to the surface. According to the surface selection rule proposed by Creighton [78] and by Moskovits and Suh [90], the C–H stretching vibration should be relatively more enhanced when the C–H bond is perpendicular to the surface plane than when it lies parallel to the surface.

According to the surface selection rule [91, 92], for benzene derivatives the out-of-plane vibrations should be more enhanced than the in-plane ones when the benzene ring assumes a parallel position with respect to the surface and the opposite should occur when the ring lies perpendicular to the surface. Changes in the bandwidth of benzene ring modes have been used [88, 93, 94] to judge the presence of surface–ring π interaction for benzene derivatives adsorbed on silver or gold.

SERS spectra have been reported for a large number of substituted benzenes [83, 95] and other aromatic compounds. SERS spectra of benzoic acid adsorbed on silver colloids [89, 96] and at silver electrodes [89, 97] have been studied. Carboxylic acids can bind to silver either via the oxygen lone pair electrons or the carboxylate π system [98], thereby assuming different orientations on the silver surface. It was suggested [89] that the benzoate species adopted a flat or tilted orientation on the silver colloid surface but assumed a perpendicular orientation with respect to a silver electrode surface. SERS studies on 3-aminosalicylic and 2-mercaptonicotinic acid in silver colloids [82] show examples of flat and perpendicular adsorption, respectively. The SERS of these two compounds [82] show how the Raman spectra can be affected when the molecules adsorb with the aromatic ring either parallel or perpendicular to a silver surface (Figs. 2 and 3). For 3-aminosalicylic acid a flat orientation on the silver surface was suggested [82], based on the following evidence in the SERS spectrum (as shown in Fig. 2): (i) the out-of-plane bending mode $\gamma_s COO^-$, 727 cm^{-1}, was relatively much more enhanced than the symmetric stretching mode $\nu_s COO^-$, 1392 cm^{-1}, suggesting that the carboxylate group should be flat on the surface instead of perpendicular; (ii) the band assigned to the amine asymmetric bend $\delta_{as}NH_2$, 1120 cm^{-1}, was quite intense and broad, suggesting that the amine group interacts directly with the surface; (iii) the absence of the C–H stretching vibrations; (iv) the strong increase in the bandwidth for the benzene ring vibrations; (v) the presence of two strongly enhanced bands for out-of-plane ring modes γCCCH, 894 cm^{-1}, and τCCCC, 438 cm^{-1}.

Evidence was obtained for the perpendicular adsorption of 2-mercaptonicotinic acid on silver from the SERS spectrum [82], in particular from its similarities with the conventional Raman spectrum of the silver complex (as shown

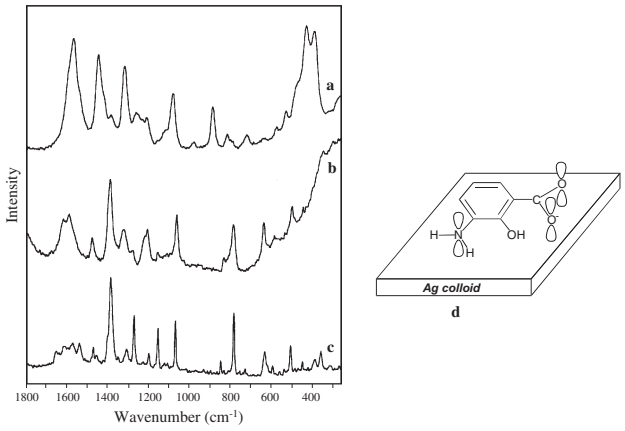


Figure 2. (a) SERS spectrum of 5×10^{-4} M 3-aminosalicylic acid solution in aqueous silver–citrate colloid, pH 6.0; (b) conventional Raman spectrum of 3-aminosalicylic acid in 3M aqueous NaOH solution; (c) conventional Raman spectrum of solid 3-aminosalicylic acid; (d) flat adsorption of 3-aminosalicylic acid on silver. Reprinted with permission from [82], H. I. S. Nogueira, *Spectrochim. Acta A* 54, 1461 (1998). © 1998, Elsevier Science.

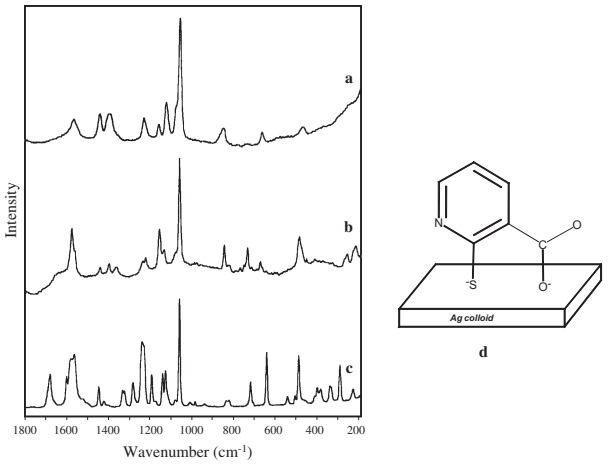


Figure 3. (a) SERS spectrum of 5×10^{-4} M 2-mercaptonicotinic acid solution in aqueous silver–citrate colloid, pH 6.0; (b) conventional Raman spectrum of solid silver complex of 2-mercaptonicotinic acid; (c) conventional Raman spectrum of solid 2-mercaptonicotinic acid; (d) perpendicular adsorption of 2-mercaptonicotinic acid on silver. Reprinted with permission from [82], H. I. S. Nogueira, *Spectrochim. Acta A* 54, 1461 (1998). © 1998, Elsevier Science.

in Fig. 3): (i) the carboxylate group vibrations for the in-plane $\nu_s COO^-$ stretch at 1393 cm^{-1} and δCOO^- bend at 849 cm^{-1} were seen in the SERS spectrum as strong and broad bands; the relatively much stronger enhancement of $\nu_s COO^-$ suggested that this group was in a perpendicular position to the silver surface; (ii) the in-plane ring vibrations ν_{as}CX at 1123 cm^{-1}, νCC at 1078 cm^{-1}, and δCC at 664 cm^{-1}, involving also the stretching of the C–S band, were stronger in the SERS spectrum; (iii) the absence of the out-of-plane τCCCC, seen at 733 cm^{-1} in the conventional Raman spectrum of the silver complex, and the relatively weaker enhancement in the SERS spectrum of τCCCC at 470 cm^{-1}; (iv) the C–H stretching was seen as a strong peak at 3068 cm^{-1}; (v) all the other enhanced modes in the SERS spectrum were in-plane modes, consistent with a perpendicular position to the silver surface; (vi) the shoulder at 245 cm^{-1} in the SERS spectrum was assigned to the Ag–S stretch, seen in the spectrum of the Ag complex as a stronger band at 257 cm^{-1}. Comparisons of the conventional Raman spectra of metal complexes (in particular of the silver complex) with the SERS spectra of the ligands have shown that there is a close relationship between complexation and adsorption (in particular for the case of a perpendicular orientation on the metal surface), as reported for 1,8-naphthyridine [99], *o*-mercaptobenzoic acid [100], and 1,2-benzenedithiol [88].

SERS studies on other substituted benzenes include salicylic acid [101], anthranilic acid [102], *p*-aminobenzoic acid [103, 104], 2-aminophenol [83], 1,2- and 1,4-benzenedithiol [88, 105]; many SERS studies were also reported on other aromatic compounds such as 1,1′-bi-2-naphthol [106], 4,4′-dicyanobiphenyl [107], quinoline derivatives [108], pyridine carboxylic acids [109], and other pyridine derivatives [110]. The first SERS reports were made on the study of pyridine; the report made by Fleischmann et al. [111] in 1974, on the Raman spectrum of a monolayer of pyridine adsorbed on a silver electrode surface, was actually the first SERS observation.

3.2.2. SERS as a Tool for Trace Detection

The strong enhancement of the normally weak Raman signals makes SERS an interesting technique for detecting low concentration samples or unknown low level contaminants in a sample [104, 112, 113]. SERS reports are found in the literature related to the detection of several organic compounds with specific interest, such as drugs and biological materials (the anti-HIV drug AZT [114], amphetamine drugs [115], sulphur- and nitrogen-containing drugs [116, 117]); detection and analysis for monitoring water contamination [118]; detection and identification of very low gas concentrations of substances that are adsorbed irreversibly at silver surfaces [119]; detection of trace amounts of atmospheric contaminants in aerosols [120]; detection of trace organic pollutants in aqueous systems [121]; and spectral fingerprinting of stereoisomer drugs at trace concentration levels [122]. The use of colloids in nonaqueous solvents allows an expansion of the organic compounds that may be detected by SERS [123].

The potential of colloidal silver and gold sols for the purpose of semiquantitative trace analysis of organic molecules by SERS has been investigated. The value of colloidal SERS for analytical purposes, in particular when employing silver sols, is described in one report to be limited to qualitative analysis [124]. Another study [125] has, for example, shown that the dye nuclear fast red can be detected and determined semiquantitatively by SERS in the concentration range of 10^{-8} to 10^{-6} M. Though the use of SERS for quantitative analysis is controversial, it has been demonstrated that it is a good technique for trace detection. SERS is important mostly with respect to the low limits of detection that can be obtained (e.g., 10 ppb for the sensitivity limit of cyanide detection in water [118], or 10^{-8} M for the dye nuclear fast red [125]).

Surface-enhanced resonance-Raman scattering (SERRS) combines two previously known effects, namely resonance-Raman scattering and SERS. The enhancement in SERRS is very much greater than either resonance-Raman or SERS enabling lower detection limits. SERRS is obtained by using a molecule with a chromophore as the adsorbate and selecting the excitation frequency to match or nearly match that of the chromophore. With SERRS, sensitivities down to 6.87×10^{-18} M were obtained using rhodamine dye [112].

4. SERS STUDIES USING NANOPARTICLES AND ITS ASSEMBLIES

Nanotechnology aims to fabricate functional devices in which their size decreases down to the nanometer range and molecular dimensions. In particular, scanning probe microscopies provide the tools to investigate the properties of functional nanodevices. More recently, new developments have been achieved with the use of Raman spectroscopy to probe surface properties with molecular resolution. It is particularly interesting to explore these findings in correlation with measurements of electron transport properties through

molecular frameworks, or for optical detection using single molecules. SERS has been generally performed on metal colloids; however, even when nanoparticles with low standard deviations are used, the spectrum of single particles has to be obtained to get optical information as a function of particle size. The use of Raman spectroscopy coupled with the use of well defined metal nanocrystals is a valuable tool which complements the single molecule detection performed by other techniques such as scanning tunneling microscopy, fluorescence measurements, or electrochemical methodologies.

4.1. SERS Substrates Based on Metal Nanoparticles

The preparation of substrates for SERS has many variations, but essentially a very finely divided sample of the metal is used either pure or in a support. The most common substrates are silver and gold colloids, but other sorts of substrates have been developed. SERS has been observed on, for example, cellulose acetate films doped with Ag fine particles [126], silver hydrosols on filter paper [120], or silver-coated alumina/glass substrates [127]. Figure 4 shows a diagram of a few examples of substrate types that can be used in SERS studies. Modification of the substrates can also be achieved by chemical coatings that can be either passive or active [119]. The passive coatings simply provide an affinity for the analyte and can act as an internal standard for normalizing the signal. Active coatings exhibit a different Raman spectrum in the presence of the analyte.

Research on metal nanoparticles led to its use as SERS substrates, in particular for silver and gold. Metal particle size and shape and the possibility of formation of nanoparticulate aggregates contribute to distinct characteristics as SERS substrates. Also, the development of a variety of solid surface-based probes based on metallic nanostructures gave a high degree of reproducibility to SERS that could not be achieved with the use of colloids [128]. The strict control of surface parameters is possible through the use of solid materials yielding the metallic nanostructures. Several types of metallic nanostructures have been developed (Fig. 4) such as metal nanoparticle island films, metal-coated polymer nanospheres, or metal-coated nanoparticles (of alumina, titanium oxide, and other oxides).

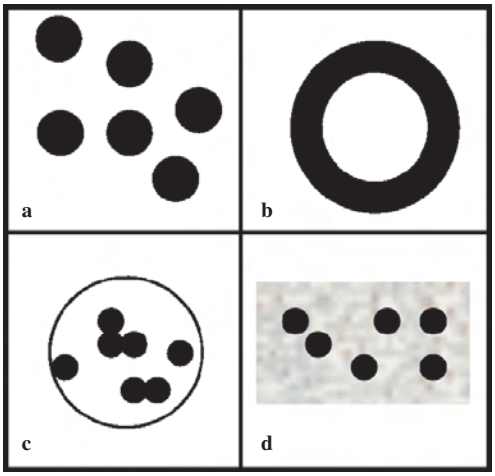


Figure 4. Diagram representation of substrates containing SERS active metal nanoparticles. (a) SERS active metal nanoparticles; (b) metal coated particles; (c) metal nanoparticles clustered into islands; (d) metal nanoparticles dispersed in a polymer substrate.

Nanostructured materials have recently been prepared by mixing silica and silver. Stable substrates were obtained by impregnating silica with silver in water followed by calcination [129], or by deposition [130] or evaporation [131] of silver over silica nanospheres. In the preparation of SERS substrates by the impregnation of silica with silver [129], the solid silica support was applied not only to disperse the metal but also to stabilize the metallic phase (by preserving the aggregation state of the Ag particles) making the SERS activity stable. A silver particle size of 6–16 nm was found to be effective for SERS. In another study [132], SERS substrates were prepared by photoreduction of silver ions onto silica colloidal nanoparticles. Silver photoreduction was achieved by irradiation with a 514.5 nm laser; the same laser radiation was employed in the SERS measurements. One advantage of the colloidal dispersions over the solid substrates is the resistance to damage caused by laser irradiation, due to the continuous agitation of the colloidal particles in the aqueous environment. Other potential SERS substrates were obtained using hybrid particles in colloidal suspensions, such as Ag–CdS [133], Ag–ZnO [134], Ag–TiO_2 [135, 136], and Ag–SiO_2 [137]. Core–shell nanoparticles have been tested as SERS substrates, using both metallic core and shell [138], or a dielectric core and a metallic shell [139]. Au/Ag core–shell nanoparticles were prepared by growing an Ag shell onto a 12 nm Au core; the SERS activity was investigated by using 2,4-dimethylpyridine as the probe [138]. A series of silica core–silver shell nanoparticles were constructed [139] using 65 and 79 nm cores, upon which silver layers ranging from 5 to 20 nm were deposited. Systematic variation of the internal geometry of the silica core–silver shell nanoparticle allowed the local electromagnetic field at the nanoparticle surface to be precisely controlled, as monitored by SERS. The SERS enhancement appears to be directly and exclusively due to nanoparticle geometry [139].

A solution method to obtain silver nanostructured thin films was reported [140], using a silver complex as the starting compound. This method consisted of the reduction of a silver complex to obtain the silver metal particles. The coordinated ligand in the starting silver complex (e.g., 2,2′-dithiodipyridine, 2dtpy, in the complex [Ag(2dtpy)]NO_3) functions as a capping agent to the Ag particle surface, avoiding the agglomeration of the nanoparticles formed, and, simultaneously, may act as a linker to the Ag nanoparticles, hence promoting the formation of a three-dimensional network. Compared to the thin film deposition of metal nanoparticles formed from conventional reductive techniques, this method offers the potential to design metal nanostructures using the coordinating organic ligand as an active template. In this sense, this preparative approach is strongly dependent on the adsorption modes of the organic ligand at the metal nanoparticle surface. SERS studies were performed [140] to investigate the adsorption of the 2dtpy ligand on the Ag nanoparticles that constitute the film. The Raman spectra of the film Ag/2,2′-dithiodipyridine, the starting solid Ag(I) complex, [Ag(2dtpy)]NO_3, and the solid

2dtpy ligand are shown in Figure 5, together with the SERS spectrum of 2dtpy in a silver aqueous colloid. The Raman spectrum of the film Ag/2,2′-dithiodipyridine (Fig. 5b) is remarkably similar to the SERS spectrum of 2dtpy (Fig. 5a) in the Ag colloid. This similarity indicates that the interaction mode of 2dtpy with silver in both cases should be very close with respect to both the atoms bound to silver and the adsorption position of the 2dtpy molecules. The transmission electron microscope images of the film showed spheroidal Ag nanocrystals (typically less than 15 nm diameter) evenly distributed, in agreement with the formation of an hybrid inorganic/organic structure of interacting Ag/2dtpy clusters.

Silver and gold nanoparticle monolayers deposited on silanized glass surfaces, have been used as SERS active substrates [141, 142]. The localized surface plasmon resonance on such arrays can be tuned to a variety of wavelengths by varying the particle shape, size, and spacing. Tunable localized surface plasmon resonance of silver nanoparticle arrays has also been achieved by nanosphere lithography [143]. The wavelength corresponding to the extinction maximum of the localized surface plasmon resonance could be systematically tuned from 300 to 6000 nm, by using precise lithographic control of nanoparticle size, height, and shape, and dielectric encapsulation of the nanoparticles in SiO_x [143].

Nanostructured thin-film materials have been prepared [144] by using size-fractionated colloidal silver nanoparticles, in the size ranges of 30–50, 50–80, 80–100, and >100 nm. Figure 6 presents an atomic force microscope (AFM) image of a nanostructured thin film prepared from 50–80 nm silver colloidal particles, showing the presence of distinct nanoparticles in an uneven distribution. These nanostructured films provide a reproducible SERS substrate for ultrasensitive detection and spectroscopy of native biological molecules [144]. As shown in Figure 7, the SERS signal intensities using this type of substrate are 5 to 10 times higher than those obtained with alternative methods [144]. Nanoparticles have already been used together with SERS, for DNA and RNA detection [145]. Several recent SERS studies on molecules of biological interest using metal nanoparticles as substrates, can be found [146–149] including SERS in single living cells [150].

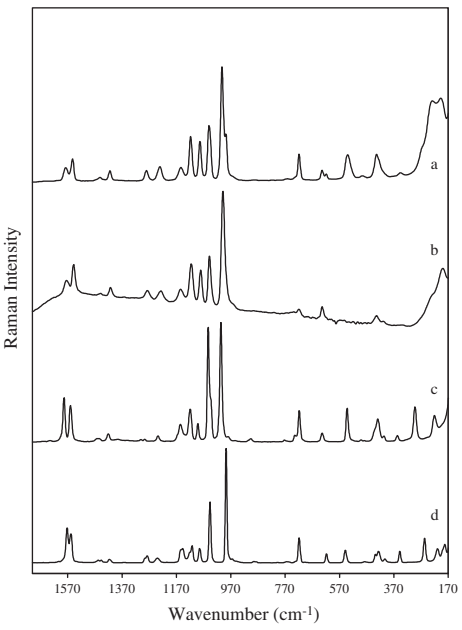


Figure 5. SERS spectrum of 2dtpy ligand in a silver aqueous colloid (a) and Fourier transform Raman spectra of the solids: Ag/2dtpy film (b), [Ag(2dtpy)]NO_3 complex (c), and 2dtpy ligand (d). Reprinted with permission from [140], H. I. S. Nogueira et al., *J. Mater. Chem.* 12, 2339 (2002). © 2002, The Royal Society of Chemistry.

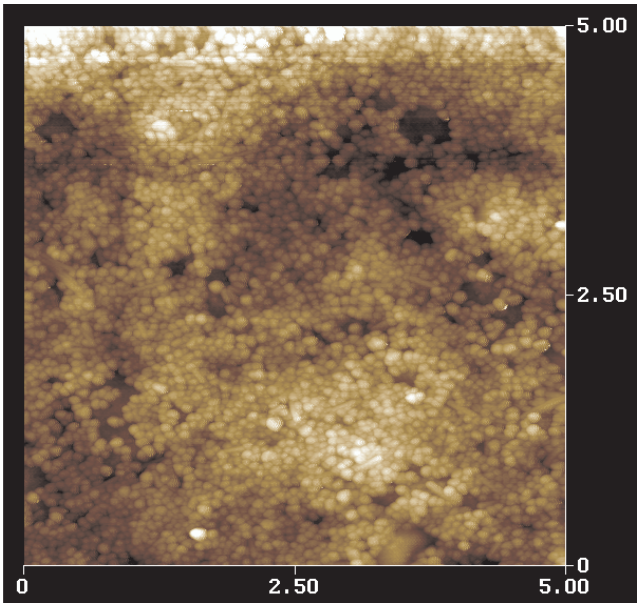


Figure 6. Tapping-mode AFM image of a nanostructured thin film prepared from 50–80 nm Ag colloidal particles. Reprinted with permission from [144], D. J. Maxwell et al., *Chem. Mater.* 13, 1082 (2001). © 2001, American Chemical Society.

4.2. Probing Single Molecules Coupled to Metal Nanoparticles Using SERS

Ultrasensitive Raman spectroscopy at the single-molecule level has been achieved in the last five years [75, 80, 81, 151–153] by exploiting the SERS effect together with the large field enhancements obtained with single nanoparticles or aggregates of nanosized metal particles. The extremely large SERS enhancement [75] seems to be a very local effect with high spatial confinement, and the single analyte has to find a special hot area (for strong field enhancement) or a hot site (for electronic enhancement). Two methods have been used to achieve single-molecule sensitivity in a SERS experiment: nonresonant near-infrared experiments on single molecule attached to silver or gold colloid nanoparticle aggregates in solution [81, 154, 155] and resonance studies performed on fixed single silver nanoparticles [80, 156]. In the nonresonant method, extremely low concentrations (on the order of 10^{-12} to 10^{-14} M) and small probed volumes (on the order of femtoliters to picoliters) ensure that only one target molecule, on average, can be probed at one time; enhancement factors on the order of 10^{14} were obtained [81]. Figure 8 shows a scheme of a typical single-molecule

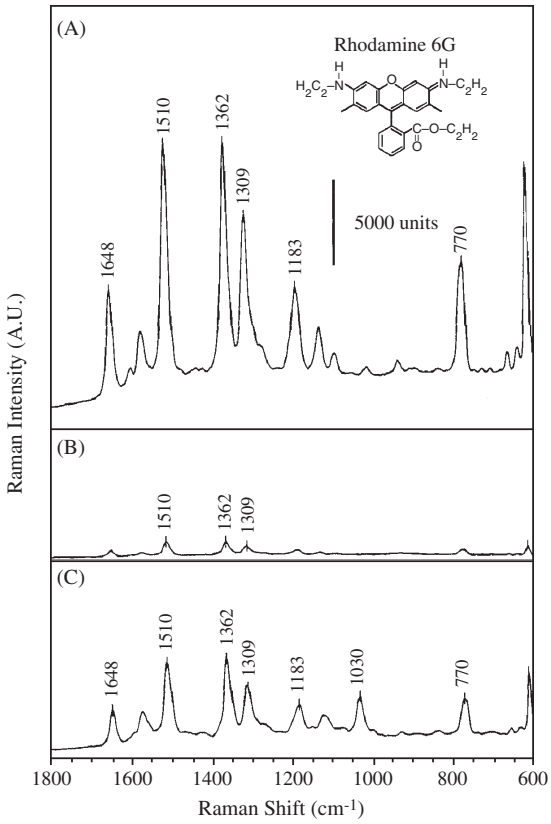


Figure 7. Comparison of SERS intensities between nanostructured thin films and colloidal nanoparticles immobilized on chemically modified surfaces. (A) Film prepared from fractionated Ag colloids (>100 nm fraction); (B) unfractionated colloids immobilized on a glass slide surface by using mercaptopropyl trimetroxysilane; and (C) unfractionated colloids immobilized on a glass surface by using polylysine. Laser wavelength = 633 nm; excitation intensity = 3 mW; data integration time = 1 s; and R6G concentration = 1×10^{-7} M. Reprinted with permission from [144], D. J. Maxwell et al., *Chem. Mater.* 13, 1082 (2001). © 2001, American Chemical Society.

SERS experiment performed in silver or gold colloidal solution [75, 149]. Spectra were excited by an argon-ion pumped continuous wave Ti:sapphire laser operating at 830 nm with a power of about 100–200 mW at the sample. A microscope attachment was used for laser excitation and collection of the Raman scattered light. The analyte was provided as a solution at concentrations smaller 10^{-11} M, which was added to the solution of small colloidal clusters [75].

In the single-particle SERS studies [80], nanometer-sized silver particles were used to amplify the spectroscopic signatures of adsorbed single molecules enormously. Simultaneously, the size-dependent properties of the nanostructures could be examined at the single-particle level. A silver colloid with adsorbed rhodamine 6G molecules was used [80], immobilized on polylysine coated glass surfaces (average silver particle size of about 35 nm). A very small number of Ag nanoparticles, called "hot particles," exhibited unusually high enhancement efficiencies. For single rhodamine 6G molecules adsorbed on the selected nanoparticles, the intrinsic Raman enhancement factors were on the order of 10^{14} to 10^{15}. Based on these findings, a SERS screening and enrichment method was reported for exploring the size and shape

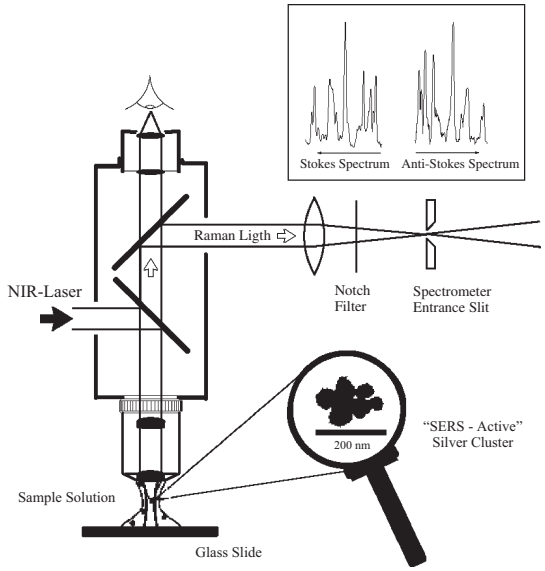


Figure 8. Schematic diagram of a single-molecule SERS experiment; a microscope is used for laser excitation and collection of the Raman scattered light. The inset shows an electron micrograph of typical SERS-active colloidal clusters. Reprinted with permission from [149], K. Kneipp et al., *J. Phys.: Condens. Matter* 14, R579 (2002). © 2002, Institute of Physics Publishing.

diversities of nanometer-scale colloidal particles [19]. The described highly active nanoparticles can be used in single-molecule spectroscopy, ultrasensitive chemical analysis, and nanomaterials design [144].

4.3. SERS as a Probe to the Chemical Nature of the Nanoparticle Surface

There is a complementary effect on SERS studies in the sense that it can be focused on the study of the internal modes of the adsorbates or on the adsorbate–metal surface modes, providing information on the type of surfaces of a specific nanometal. The correlation between the optical enhancement with the morphological characteristics of metal nanoparticles makes SERS a potential technique for nanomaterials design [80]. In this sense, standard molecular adsorbates can be used to discriminate specific particles sizes of nanoparticles assemblies.

In this context, there is interest to extend SERS studies to a variety of metal surfaces. This is particularly interesting when associated with the study of catalytic processes occurring at various metal surfaces. A possible approach which has been investigated involves the use of ultrathin metals coating a SERS active substrate (Ag or Au) [157]. Although in these cases valuable information about the metal surfaces can be obtained, there is also the possibility to have enhancement of the Raman signal from the adsorbate in contact with the active SERS substrate or at the metal/metal interface. The ideal situation, and also the most difficult to obtain, is to get the SERS spectrum on a wide range of bare metal surfaces (i.e., in a situation analogous to the SERS obtained when Ag and Au substrates are used). This has attracted considerable attention since the observation of the C–O stretching band from CO adsorbed

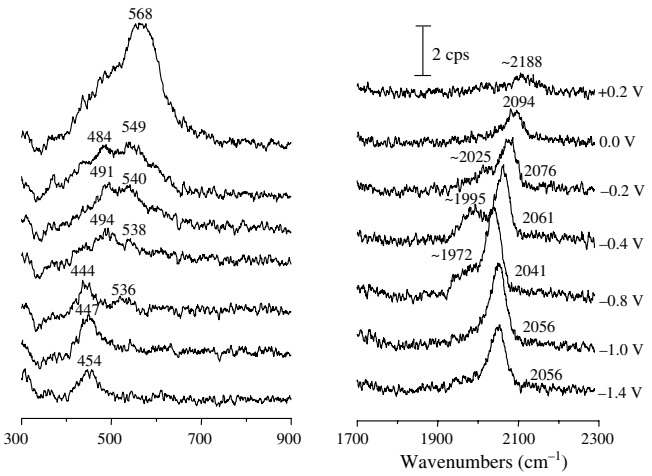


Figure 9. Potential-dependent surface Raman spectra from CO adsorbed irreversibly at a roughened Pt electrode. Reprinted with permission from [157], Z. Q. Tian et al., *J. Phys. Chem. B* 101, 1338 (1997). © 1997, American Chemical Society.

on a platinized Pt electrode [158]. Since then considerable efforts have been made to obtain SER spectra using bare Pt electrodes, although low signal intensities were generally observed [157]. Some of these limitations will tend to be overcome with advances in Raman instrumentation. An elucidative example of this was the use of confocal microprobe Raman spectroscopy and a electrochemical pretreatment of Pt electrodes to extend the detailed surface Raman studies to bare Pt substrates, using SCN^- and CO as the adsorbates (Fig. 9) [157].

Another approach to extend SERS studies to other metal surfaces involves the nanoengineering of specific metal substrates. For example, metal nanorod arrays with high SERS activity have been used as a diagnostic tool to metal nanowires with specific cross sections [159]. By analysis of SERS spectra, namely the band frequency shifts and the intensity of specific probe molecules (e.g., pyridine, SCN^-) adsorbed onto metal surfaces, one may conclude with special electronic properties of the metal nanostructures [159]. A major drawback that has limited SERS as a tool for nanometal design is that the high Raman enhancement has been limited to silver, gold, and, to a lesser extent, to copper. It should be noted that in the latter case [159], due to the nanorod morphology, SERS activity was observed even for metals which usually are non-SERS active such as in the case of several transition metals, including nickel and cobalt. In this case, two-dimensional arrays of metal nanowires (15–70 nm) were prepared by a template synthesis method involving the electrodeposition of the respective metal into nanoholes of the anodic aluminum oxide, followed by removal of the film. SERS on net transition metals has been studied [160] (e.g., Pt, Ru, Rh, Pd, Fe, Co, Ni, and their alloys) by developing various roughening procedures. An approach that replaces the randomly roughened surface with ordered nanorod arrays of transition metals was introduced as a promising class of highly SERS-active substrates [160].

5. PERSPECTIVES

The chemical control of the properties of metal nanostructures associated with recent developments on Raman spectroscopic methods has increased the possibility to apply this technique in innovative contexts. For example, the extension of SERS studies to other metal surfaces rather than Ag and Au, by tailoring specific metal nanostructures, will provide new insights about the mechanisms behind the SERS effect. The study of size-dependent properties of nanometals by SERS screening at the single-molecule level will provide novel knowledge on the understanding of the mechanisms behind surface modification and self-assembly methods. Using single-molecule SERS, the size-dependent properties of the nanostructures can be examined at the single-particle level, and, simultaneously, the spectroscopic features of adsorbed single molecules are enormously amplified. The combination of these complementary effects will give an increasing importance of single-molecule SERS for future research in the fabrication of nanodevices containing nanometals and molecular units.

GLOSSARY

Metal nanoclusters Nanoparticles composed of up to hundreds of atoms, with dimensions typically less than 3 nm in diameter, which are stabilized with capping organic ligands.

Metal nanocrystals Large nanoparticles which may have dimensions superior to 100 nm, constituents of most of the metal colloids employed in SERS. In metal nanocrystals the internal crystalline structure is similar to the macrocrystalline lattice but it shows distinct physical and chemical properties from the bulk materials and from nanoclusters of metallic atoms.

Raman spectrum When monochromatic light of frequency ν_0 is incident on a transparent sample and the scattered light is dispersed by a monochromator and subsequently measured in intensity, the spectrum of the inelastically scattered radiation (the Raman spectrum of the sample) is obtained.

Single-molecule SERS SERS measured on a single molecule using single-metal nanoparticles or aggregates of nanosized metal particles, giving stronger enhancement of the Raman signal up to 10^{14}. The extremely large SERS enhancement seems to be a very local effect with high spacial confinement, and the single analyte has to find a special hot area (for strong field enhancement) or a hot site (for electronic enhancement).

Surface enhanced Raman scattering (SERS) Refers to the observation that for certain molecules adsorbed on specially prepared metal surfaces a Raman spectrum is obtained whose intensity is enhanced by a factor of up to 10^6. The large enhancement factors in SERS are obtained through a combination of an electromagnetic effect and chemical interactions between the adsorbate molecule and the surface.

Surface modification Derivatization of the metal particle surface using either an inorganic phase, an organic ligand, or both. The derivatization process should not disturb the integrity of the metal core.

REFERENCES

1. M. Faraday, Relations of gold and other metals to light, in "Experimental Researches in Chemistry and Physics," p. 391. Taylor and Francis, London, 1859.
2. G. Mie, *Ann. Phys.* 25, 377 (1908).
3. D. L. Jeanmaire and R. P. Van Duyne, *J. Electroanal. Chem.* 84, 1 (1977).
4. M. G. Albrecht and J. A. Creighton, *J. Am. Chem. Soc.* 99, 5215 (1977).
5. U. Kreibig, H. Bönnemann, and J. Hormes, Nanostructured metal clusters and colloids, in "Handbook of Surfaces and Interfaces of Materials" (H. S. Nalwa, Ed.), Vol. 3. Academic Press, San Diego, 2001.
6. G. Schmid, *Polyhedron* 7, 2321 (1988).
7. L. J. Jongh, Electronic properties of metalcluster compounds: Nanophase materials from chemical synthesis, in "Nanophase Materials—Synthesis, Properties, Applications" (G. C. Hadjipanayis and R. W. Siegel, Eds.), NATO ASI Series, Vol. 260, p. 349. Kluwer Academic, Dordrecht, 1994.
8. C. N. R. Rao, G. U. Kulkarni, A. Govindaraj, B. C. Satishkumar, and P. J. Thomas, *Pure Appl. Chem.* 72, 21 (2000).
9. M. P. Pileni, *Pure Appl. Chem.* 72, 53 (2000).
10. S. N. Khanna, Effect on properties of reduced size and dimensions, in "Handbook of Nanoscale Materials" (A. N. Goldstein, Ed.), p. 1. Dekker, New York, 1997.
11. U. Kreibig and M. Vollmer, "Optical Properties of Metal Clusters," Springer Series in Materials Science Vol. 25. Springer-Verlag, Heidelberg, 1995.
12. J. M. Thomas and W. J. Thomas, "Principles and Practice of Heterogeneous Catalysis." VCH, Weinheim, 1997.
13. J. Pedersen, S. Bjørnholm, J. Borggreen, K. Hansen, T. P. Martin, and H. D. Rasmussen, *Nature* 353, 733 1991.
14. R. L. Whetten, J. T. Khoury, M. M. Alvarez, S. Murthy, I. Vezmar, Z. L. Wang, P. W. Stephens, C. L. Cleveland, W. D. Luedtke, and U. Landman, *Adv. Mater.* 8, 428 (1996).
15. C. F. Bohren and D. R. Huffman, "Absorption and Scattering of Light by Small Particles." Wiley, New York, 1983.
16. R. H. Doremus, *J. Chem. Phys.* 40, 2389 (1964).
17. J. Turkevich, G. Garton, and P. C. Stevenson, *J. Colloid. Sci.* 9, 26 (1954).
18. P. C. Lee and D. J. Meisel, *J. Phys. Chem.* 86, 3391 (1982).
19. S. R. Emory and S. Nie, *J. Phys. Chem. B* 102, 493 (1998).
20. F. Klim, J. H. Song, and P. Yang, *J. Am. Chem. Soc.* 124, 14316 (2002).
21. Y. Wang and N. Toshima, *J. Phys. Chem. B* 101, 5301 (1997).
22. F. Fievet, J. P. Lagier, B. Blin, B. Beaudoin, and M. Figlarz, *Solid State Ionics* 32/33, 198 (1989).
23. I. Pastoriza-Santos and L. M. Liz-Marzán, *Pure Appl. Chem.* 72, 83 (2000).
24. R. L. Garrel and R. H. Schultz, *J. Colloid Interface Sci.* 105, 483 (1985).
25. M. P. Pileni, *J. Phys. Chem.* 97, 6961 (1993).
26. F. C. Meldrum, N. A. Kotov, and J. H. Fendler, *J. Chem. Soc. Faraday Trans.* 90, 673 (1994).
27. J. Tanori and M. P. Pileni, *Adv. Mater.* 7, 862 (1995).
28. M. P. Pileni, J. Tanori, and A. Filankembo, *Colloids Surfaces A* 123–124, 561 (1997).
29. J. B. Nagy, D. Barette, A. Fonseca, L. Jeunieau, Ph. Monnoyer, P. Piedigrosso, I. Ravet-Bodart, J. P. Verfaillie, and A. Wathelet, Nanoparticles in microemulsions: A general approach, in "Nanoparticles in Solids and Solutions" (J. H. Fendler and I. Dékány, Eds.), NATO ASI Series, Vol. 71 (1996).
30. A. Taleb, C. Petit, and M. P. Pileni, *Chem. Mater.* 9, 950 (1997).
31. M. Y. Han, C. H. Quek, W. Huang, C. H. Chew, and L. M. Gan, *Chem. Mater.* 11, 1144 (1999).
32. K. Torigoe and K. Esumi, *Langmuir* 11, 4199 (1995).
33. A. Szücs, Z. Király, F. Berger, and I. Dékány, *Colloids Surfaces A* 139, 109 (1998).
34. Z. Király, I. Dékány, A. Mastalir, and M. Bartók, *J. Catal.* 161, 401 (1996).
35. G. Rupprechter, K. Hayek, and H. Hofmeister, *J. Catal.* 173, 409 (1998).
36. L. Tosheva and J. Sterte, *Chem. Commun.* 1112 (2001).
37. W. Yu and H. Liu, *Chem. Mater.* 10, 1205 (1998).
38. N. A. Dhas, C. P. Raj, and A. Gedanken, *Chem. Mater.* 10, 1446 (1998).
39. J. S. Bradley, E. W. Hill, C. Klein, B. Chaudret, and A. Duteil, *Chem. Mater.* 5, 254 (1993).
40. M. Reetz and W. Helbig, *J. Am. Chem. Soc.* 116, 7401 (1994).
41. P. Mulvaney, L. M. Liz-Marzán, M. Giersig, and T. Ung, *J. Mater. Chem.* 10, 1259 (2000).
42. M. Brust, M. Walker, D. Bethell, D. J. Schiffrin, and R. Whyman, *J. Chem. Soc. Commun.* 801 (1994).
43. M. Brust, J. Fink, D. Bethell, D. J. Schiffrin, and C. Kieley, *J. Chem. Soc. Chem. Commun.* 1655 (1995).
44. R. L. Whetten, M. N. Shafigullin, J. T. Khoury, T. G. Schaaff, I. Vezmar, M. M. Alvarez, and A. Wilkinson, *Acc. Chem. Res.* 32, 397 (1999).
45. M. J. Hostetler, S. J. Green, J. J. Stokes, and R. W. Murray, *J. Am. Chem. Soc.* 118, 4212 (1996).
46. R. P. Andres, J. D. Bielefeld, J. I. Henderson, D. B. Janes, V. R. Kolagunta, C. P. Kubiak, W. Mahoney, and R. G. Osifchin, *Science* 273, 1690 (1996).
47. K. V. Sarathy, G. U. Kulkarni, and C. N. R. Rao, *Chem. Commun.* 537 (1997).
48. W. Li, J. A. Virtanen, and R. G. Penner, *J. Phys. Chem.* 98, 11751 (1994).
49. W. Li, J. A. Virtanen, and R. G. Penner, *Langmuir* 11, 4361 (1995).
50. D. V. Leff, L. Brandt, and J. R. Heath, *Langmuir* 12, 4723 (1996).
51. T. Yonezawa, H. Matsune, and T. Kunitake, *Chem. Mater.* 11, 33 (1999).
52. K. V. Sarathy, P. J. Thomas, G. U. Kulkarni, and C. N. R. Rao, *J. Phys. Chem. B* 103, 399 (1999).
53. B. A. Korgel, S. Fullam, S. Connolly, and D. Fitzmaurice, *J. Phys. Chem. B* 102, 8379 (1998).
54. L. Han, J. Luo, N. N. Kariuki, M. M. Maye, V. W. Jones, and C. J. Zhong, *Chem. Mater.* 15, 29 (2003).
55. E. U. Thoden van Velzen, J. f. J. Engbersen, and D. N. Reinhoudt, *J. Am. Chem. Soc.* 116, 3597 (1994).
56. H. Adams, F. Davis, and C. J. M. Stirling, *J. Chem. Soc., Chem. Commun.* 2527 (1994).
57. A. N. Sipway, M. Lahav, R. Blonder, and I. Willner, *Chem. Mater.* 11, 13 (1999).
58. B. I. Ipe, K. G. Thomas, S. Barazzouk, S. Hotchandani, and P. V. Kamat, *J. Phys. Chem. B* 106, 18 (2002).
59. S. L. Tripp, S. V. Pusztay, A. E. Ribbe, and A. Wei, *J. Am. Chem. Soc.* 124, 7914 (2002).
60. C. V. Raman and K. S. Krishnan, *Nature* 121, 501 (1928).
61. C. V. Raman, *Nature* 121, 619 (1928).
62. C. V. Raman and K. S. Krishnan, *Nature* 121, 711 (1928).
63. D. R. Porterfield and A. Campion, *J. Am. Chem. Soc.* 110, 408 (1988).
64. G. F. Puppels, F. F. de Mul, C. Otto, J. Greve, M. Robert-Nicoud, D. J. Arndt-Jovin, and T. M. Jovin, *Nature* 347, 301 (1990).
65. D. A. Long, "Raman Spectroscopy," p. 139. McGraw–Hill, London, 1977.
66. J. E. Pemberton, R. L. Sobocinski, M. A. Bryant, and D. A. Carter, *Spectrosc. Int.* 2, 26 (1990) [also appeared in *Spectroscopy* 5, 26 (1990)].
67. D. L. Rousseau and P. F. Williams, *J. Chem. Phys.* 64, 3519 (1976).
68. J. M. Friedman, D. L. Rousseau, and V. E. Bondybey, *Phys. Rev. Lett.* 37, 1610 (1976).
69. P. F. Williams and D. L. Rousseau, *Phys. Rev. Lett.* 30, 951 (1973).

70. D. L. Rousseau, J. M. Friedman, and P. F. Williams, The resonance Raman effect, in "Raman Spectroscopy of Gases and Liquids" (A. Weber, Ed.), p. 203. Springer-Verlag, Berlin, 1979.
71. P. F. Williams, D. L. Rousseau, and S. H. Dworetsky, *Phys. Rev. Lett.* 32, 196 (1974).
72. J. A. Creighton, *Anal. Proc.* 30, 28 (1993).
73. "Surface Enhanced Raman Scattering" (R. K. Chang and T. E. Furtak, Eds.). Plenum, New York, 1982.
74. "Proceedings of the 18th International Conference on Raman Spectroscopy" (J. Mink, G. Jalsovszky, and G. Keresztury, Eds.). Wiley, Chichester, 2002.
75. K. Kneipp, H. Kneipp, I. Itzkan, R. R. Dasari, and M. S. Feld, *Chem. Rev.* 99, 2957 (1999).
76. M. Moskovits, *Rev. Mod. Phys.* 57, 783 (1985).
77. M. Moskovits, *J. Chem. Phys.* 69, 4159 (1978).
78. J. A. Creighton, *Surf. Sci.* 124, 209 (1983).
79. J. A. Creighton, *Surf. Sci.* 173, 665 (1986).
80. S. Nie and S. R. Emory, *Science* 257, 1102 (1997).
81. K. Kneipp, Y. Wang, H. Kneipp, L. T. Perelman, I. Itzkan, R. R. Dasari, and M. S. Feld, *Phys. Rev. Lett.* 78, 1667 (1997).
82. H. I. S. Nogueira, *Spectrochim. Acta A* 54, 1461 (1998).
83. W. P. Griffith and T. Y. Koh, *Spectrochim. Acta A* 51, 253 (1995).
84. J. A. Creighton, *Surf. Sci.* 158, 211 (1985).
85. M. Moskovits, *J. Chem. Phys.* 77, 4408 (1982).
86. M. Moskovits and J. S. Suh, *J. Phys. Chem.* 88, 5526 (1984).
87. M. Moskovits, D. P. DiLella, and K. J. Maynard, *Langmuir* 4, 67 (1988).
88. S. H. Cho, Y. J. Lee, M. S. Kim, and K. Kim, *Vibr. Spectrosc.* 10, 261 (1996).
89. Y. J. Kwon, D. H. Son, S. J. Ahn, M. S. Kim, and K. Kim, *J. Phys. Chem.* 98, 8481 (1994).
90. M. Moskovits and J. S. Suh, *J. Am. Chem. Soc.* 108, 4711 (1986).
91. M. Takahashi, M. Fujita, and M. Ito, *Chem. Phys. Lett.* 109, 122 (1984).
92. M. Moskovits and J. S. Suh, *J. Phys. Chem.* 88, 1293 (1984).
93. S. H. Cho, H. S. Han, D. J. Jang, K. Kim, and M. S. Kim, *J. Phys. Chem.* 99, 10594 (1995).
94. X. P. Gao, J. P. Davies, and M. J. Weaver, *J. Phys. Chem.* 94, 6858 (1990).
95. T. Y. Koh, S. J. Greaves, and W. P. Griffith, *Spectrochim. Acta A* 50, 857 (1994).
96. M. Pagannone, B. Fornari, and G. Mattei, *Spectrochim. Acta A* 43, 621 (1987).
97. G. S. Shin and J. J. Kim, *Surf. Sci.* 158, 286 (1985).
98. Y. J. Kwon, S. B. Lee, M. S. Kim, and K. Kim, *J. Mol. Struct.* 318, 25 (1994).
99. W. P. Griffith and T. Y. Koh, *J. Raman Spectrosc.* 26, 1067 (1995).
100. S. B. Lee, K. Kim, and M. S. Kim, *J. Raman Spectrosc.* 22, 811 (1991).
101. E. Wentrup-Byrne, S. Sarinas, and P. M. Fredericks, *Appl. Spectrosc.* 47, 1192 (1993).
102. S. Nie, C. G. Castillo, K. L. Bergbauer, J. F. R. Kuck Jr., I. R. Nabiev, and N. T. Yu, *Appl. Spectrosc.* 44, 571 (1990).
103. J. S. Suh, D. P. DiLella, and M. Moskovits, *J. Phys. Chem.* 87, 1540 (1983).
104. H. Park, S. B. Lee, K. Kim, and M. S. Kim, *J. Phys. Chem.* 94, 7576 (1990).
105. S. H. Cho, H. S. Han, D. J. Jang, K. Kim, and M. S. Kim, *J. Phys. Chem.* 99, 10594 (1995).
106. H. I. S. Nogueira and S. M. O. Quintal, *Spectrochimica Acta A* 56, 959 (2000).
107. C. Lee, S. J. Bae, M. Gong, K. Kim, and S. Joo, *J. Raman Spectrosc.* 33, 429 (2002).
108. M. Bolboaca, W. Kiefer, and J. Popp, *J. Raman Spectrosc.* 33, 207 (2002).
109. S. M. Park, K. Kim, and M. S. Kim, *J. Mol. Struct.* 344, 195 (1995).
110. M. M. B. Pessoa and M. L. A. Temperini, *J. Raman Spectrosc.* 33, 50 (2002).
111. M. Fleischmann, P. J. Hendra, and J. McQuillan, *Chem. Phys. Lett.* 26, 163 (1974).
112. C. Rodger, W. E. Smith, G. Dent, and M. Edmonson, *J. Chem. Soc., Dalton Trans.* 791 (1996).
113. T. Vodinh, *Sensor Actuator B* 29, 183 (1995).
114. L. Rivas, S. Sanchez-Cortes, and J. V. Garcia-Ramos, *J. Raman Spectrosc.* 33, (2001).
115. R. A. Sulk, R. C. Corcoran, and K. T. Carron, *Appl. Spectrosc.* 53, 954 (1999).
116. E. L. Torres and J. D. Winefordner, *Anal. Chem.* 59, 1626 (1987).
117. R. Montes and J. J. Laserna, *Analyst* 115, 1601 (1990).
118. W. R. Premasiri, R. H. Clarke, S. Londhe, and M. E. Womble, *J. Raman Spectrosc.* 32, 919 (2001).
119. W. Hill, B. Wehling, and D. Klockow, *Appl. Spectrosc.* 53, 547 (1999).
120. M. J. Ayora, L. Ballesteros, R. Perez, A. Ruperez, and J. J. Laserna, *Anal. Chim. Acta* 355, 15 (1997).
121. N. Weissenbacher, B. Lendl, J. Frank, H. D. Wanzenbock, B. Mizaikoff, and R. Kellner, *J. Mol. Struct.* 40, 539 (1997).
122. A. Ruperez and J. J. Laserna, *Anal. Chim. Acta* 335, 87 (1996).
123. J. Clarkson, C. Campbell, B. N. Rospendowski, and W. E. Smith, *J. Raman Spectrosc.* 22, 771 (1991).
124. J. C. Cook, C. M. P. Cuypers, B. J. Kip, and R. J. Meyer, *J. Raman Spectrosc.* 24, 609 (1993).
125. I. T. Shadi, B. Z. Chowdhry, M. J. Snowden, and R. Withnall, *Anal. Chim. Acta* 450, 115 (2001).
126. Y. Imai, Y. Kurokawa, M. Hara, and M. Fukushima, *Spectrochim. Acta A* 53, 1697 (1997).
127. Y. Li and Y. Wang, *Appl. Spectrosc.* 46, 142 (1992).
128. T. Vo-Dinh and D. L. Stokes, SERS-based Raman probes, in "Handbook of Vibrational Spectroscopy" (J. M. Chalmers and P. R. Griffiths, Eds.), Vol. 2, p. 1302. Wiley, Chichester, 2002.
129. F. X. Liu, Y. Xiao, and Y. Li, *J. Raman Spectrosc.* 32, 73 (2001).
130. Y. Kobayashi, V. Salgueirino-Maceira, and L. M. Liz-Marzán, *Chem. Mater.* 13, 1630 (2001).
131. M. Litorja, C. L. Haynes, A. J. Haes, T. R. Jensen, and R. P. Van Duyne, *J. Phys. Chem. B* 105, 6907 (2001).
132. M. Muniz-Miranda, *J. Raman Spectrosc.* 33, 295 (2002).
133. I. Honma, T. Sano, and H. Komiyama, *J. Phys. Chem.* 97, 6692 (1993).
134. S. Chen and U. Nickel, *J. Chem. Soc., Faraday Trans.* 92, 1555 (1996).
135. C. Wang, C. Liu, Y. Liu, and Z. Zhang, *Appl. Surf. Sci.* 147, 52 (1999).
136. Y. Liu, C. Liu, Z. Zhang, and C. Wang, *Spectrochim. Acta A* 57, 35 (2001).
137. T. Ung, L. M. Liz-Marzan, and P. Mulvaney, *Langmuir* 14, 3740 (1998).
138. X. H. Ji, L. Y. Wang, X. T. Zhang, Y. B. Bai, T. J. Li, Z. Z. Zhi, X. G. Kong, and Y. C. Liu, *Chem. J. Chinese Univ.* 23, 2357 (2002).
139. J. B. Jackson, S. L. Westcott, L. R. Hirsch, J. L. West, and N. J. Halas, *Appl. Phys. Lett.* 82, 257 (2003).
140. H. I. S. Nogueira, P. C. R. Soares-Santos, S. M. G. Cruz, and T. Trindade, *J. Mater. Chem.* 12, 2339 (2002).
141. W. Gotschy, K. Vonmetz, A. Leitner, and F. R. Aussenegg, *Opt. Lett.* 21, 1099 (1996).
142. N. Felidj, J. Aubard, G. Levi, J. R. Krenn, M. Salerno, G. Schider, B. Lamprecht, A. Leitner, and F. R. Aussenegg, *Phys. Rev. B* 65, 75419 (2002).
143. T. R. Jensen, M. D. Malinsky, C. L. Haynes, and R. P. Van Duyne, *J. Phys. Chem. B* 104, 10549 (2000).
144. D. J. Maxwell, S. R. Emory, S. M. Nie, *Chem. Mater.* 13, 1082 (2001).
145. Y. W. C. Cao, R. Jin, and C. A. Mirkin, *Science* 297, 1536 (2002).

146. M. Ishikawa, Y. Maruyama, J. Y. Ye, and M. Futamata, *J. Lumin.* 98, 81 (2002).
147. S. Sanchez-Cortes, R. M. Berenguel, A. Madejon, and M. Perez-Mendez, *Biomacromolecules* 3, 655 (2002).
148. L. Rivas, S. Sanchez-Cortes, and J. V. Garcia-Ramos, *Phys. Chem. Chem. Phys.* 4, 1943 (2002).
149. K. Kneipp, H. Kneipp, I. Itzkan, R. R. Dasari, and M. S. Feld, *J. Phys.: Condens. Matter.* 14, R597 (2002).
150. K. Kneipp, A. S. Haka, H. Kneipp, K. Badizadegan, N. Yoshizawa, C. Boone, K. E. Shafer-Peltier, J. T. Motz, R. R. Dasari, and M. S. Feld, *Appl. Spectrosc.* 56, 150 (2002).
151. M. Moskovits, L. L. Tay, J. Yang, and T. Haslett, *Topics Appl. Phys.* 82, 215 (2002).
152. A. Otto, *J. Raman Spectrosc.* 33, 593 (2002).
153. K. A. Bosnick, J. Jiang, and L. E. Brus, *J. Phys. Chem. B* 106, 8096 (2002).
154. K. Kneipp, H. Kneipp, V. B. Kartha, R. Manoharan, G. Deinum, I. Itzkan, R. R. Dasari, and M. S. Feld, *Phys. Rev. E* 57, R6281 (1998).
155. K. Kneipp, H. Kneipp, G. Deinum, I. Itzkan, R. R. Dasari, and M. S. Feld, *Appl. Spectrosc.* 52, 175 (1998).
156. A. M. Michaels, M. Nirmal, and L. E. Brus, *J. Am. Chem. Soc.* 121, 9932 (1999).
157. Z. Q. Tian, B. Ren, and B. W. Mao, *J. Phys. Chem. B* 101, 1338 (1997).
158. R. P. Cooney, M. Fleischmann, and P. J. Hendra, *J. Chem. Soc., Chem. Commun.* 235 (1977).
159. J. L. Yao, G. P. Pan, K. H. Xue, D. Y. Wu, B. Ren, D. M. Sun, J. Tang, X. Xu, and Z. Q. Tian, *Pure Appl. Chem.* 72, 221 (2000).
160. Z. Q. Tian, B. Ren, and D. Y. Wu, *J. Phys. Chem. B* 106, 9463 (2002).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanostructured Organic Light-Emitting Diodes

Thien-Phap Nguyen

Université de Nantes, France

Gilles Horowitz

Université Denis-Diderot, Paris, France

CONTENTS

1. INTRODUCTION

Light emission from organic materials has been known for a long time. One of the most amazing examples of organic light comes from a small insect known as a firefly. It emits light by a chemical process produced by its own biochemistry and the light emission is then called *bioluminescence*. What is remarkable is the efficiency of the conversion process, which is very high, close to unity, and at this stage it is not possible for scientists to obtain a similar performance with all the techniques they have. Some fish living in deep oceans can also emit light by bioluminescence.

Light can be also generated in organic materials by other conversion processes. When the excitation source is light, the process is called *photoluminescence*. Similarly, if the origin of the excitation is chemical, thermal, or electrical, the terms *chemiluminescence*, *thermoluminescence*, or *electroluminescence* are used to designate the light emission.

In electroluminescent devices, the electrical excitation of the material is provided by an external voltage source. A current flows through the device and, at a sufficient level of supplied energy, the light emission process starts in the emitter and the device emits light to the external environment. Electroluminescence in organic materials was first done experimentally in 1963 on anthracene monocrystals by Pope et al. [1] who obtained light emission by using a high applied voltage (400 V). Such a bias was considered too high and the materials were therefore not suitable to be used for everyday-use applications. In 1987, Tang and Van Slyke [2] using thin film deposition techniques and organic materials fabricated the first light emitting diodes (LEDs), which operated at low bias. A few years later, in 1990, Burroughes et al. [3] demonstrated that a well known conjugated polymer, poly(*p*-phenylene vinylene) (PPV), could emit light in a diode configuration with an applied voltage less than 20 V. Since those first experiments, the science of polymer and organic LEDs has progressed greatly. Talented chemists, physicists, and engineers have contributed to the development of new materials and new device designs, enabling the fabrication of devices with high performance: various colors, low turn-on voltage, long lifetime, low energy consumption, and great brightness. The physics of these new electronic devices have also been intensively studied and during the last decade; many new concepts were built using the basic knowledge from classical semiconductor devices. In parallel, nanotechnology science has emerged and is advancing constantly. In the late 1960s, Moore's law predicted that the number of transistors on a chip would roughly double every 18 months and that was verified over four decades. At the beginning of 2000, we have entered in the nanosize era. Advances in manipulating nanosized materials have already allowed improvements in electronic devices such as computer data storage, batteries, solar cells, and light emitting diodes. Understanding the potential market of nanosized devices, many countries have concentrated their efforts in developing research programs in nanoscience.

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (717–739)

The conjugation of both technologies led to a new field of research: that of nanosized polymer and organic LEDs (OLEDs). In reality, the devices studied have a size ranging from several nanometers (nm) to several micrometers (μm). Classical photolithography used in computer chip industry can provide a common size of 200 nm and is still very popular in designing electronic devices. Recent progress in patterning techniques has allowed features as small as 10 nm to be obtained with high definition and ordering. Such a tiny object is not easy to handle and even if several devices could be produced, the mass production still needs further investigation to control them, especially in the design of the surface on which will be implanted the nanocomponents and the electrical connection between them.

With the development of nano-objects, the electronic properties of the materials will change from classical physics to the less familiar quantum physics. New optical and electrical properties of nanosize materials have been discovered and more are expected to be found in the near future. Study of micro- and nano-OLEDs has only been investigated for a few years, starting with composites made of a polymer matrix containing inorganic nanoparticles. These systems were first fabricated to obtain the light emission of both the electroluminescent polymer and the nanoparticles. The second generation of small OLEDs has benefitted from the progress in patterning technology to allow fabrication of ordered arrays of pixels.

This chapter is organized as follows. In the first section, we shall give the state of the art of OLEDs showing different aspects: materials, physical processes including the light emission, charge carrier transport, and degradation mechanisms. The second section deals with the nano-OLEDs where different technologies will be described. Some aspects of the physical processes in nanosize materials will be treated. In the last part, the potential applications of nano-OLEDs are discussed.

2. ORGANIC ELECTROLUMINESCENCE

The aim of this section is to present the current state of the art in organic electroluminescence. Emphasis will be given to the way the device works and to the various issues that govern its operating process.

2.1. Operating Mode

Electroluminescence (EL) is a nonthermal emission of light from a solid traversed by an electrical current. EL in organic materials was first observed in anthracene single crystals as early as 1963 [4]. The crucial issues of the process and in particular the requirement of a double injection of charge (electrons and holes) were identified two years later [5]. Today, the most generally accepted mechanism for describing organic EL does not fundamentally differ from this early analysis. The basic device for organic EL is the diode, composed of one or several organic layers sandwiched between two conducting electrodes. The sequential steps of the conversion of electricity to light in an OLED are schematized in Figure 1. At the outset of the process, there is an injection of electrons and holes from the electrodes at each side of the diode. The carriers are then driven toward the interior of the organic layer by the electric field generated by the voltage applied to the device. Recombination of charges of opposite sign occasionally occurs when these charges meet, leading to the formation of electron–hole pairs, which rapidly decay to bound excitons. At this stage, a crucial parameter is the exciton bounding energy; excitons in organic materials are generally regarded as strongly bound excited molecules, but this is still controversial, as will be detailed in the following. Light is eventually emitted when excitons deactivate to ground state. The overall efficiency of the diode results from that of each step; therefore, each of them needs to be optimized. The steps will now be analyzed in more detail.

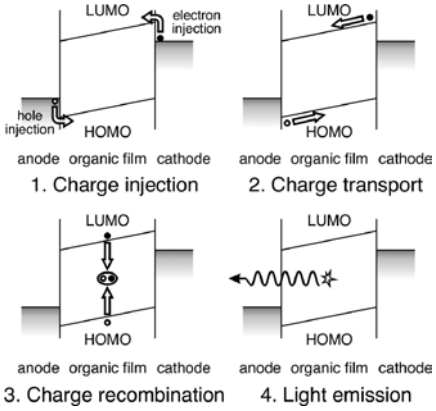


Figure 1. The four steps of the conversion of electricity to light in an organic light-emitting diode.

2.1.1. Charge Injection

Injecting charge from a highly conducting (the electrodes) into a poorly conducting (the organic layer) medium is a difficult task. In principle, charge transfer between two media occurs isoenergetically. Accordingly, there are two elementary issues in charge injection: energy level alignment and injection processes.

Energy Level Alignment This subject has been recently reviewed in detail by Ishii et al. [6]. The levels implied in charge injection from a metal into an organic layer are not of the same nature at both sides of the interface. When gathered together to from a solid, atoms and molecules see their discrete quantum energy levels widening into alternatively allowed and forbidden bands. In metals and degenerate semiconductors, the highest energy allowed band is only partially filled with electrons up to the so-called Fermi level, which constitutes therefore the level of interest in this case. The corresponding energy is the work function, Φ_m, which measures the distance between the Fermi level and the vacuum level (VL), the latter being defined as the energy of an isolated electron at rest at infinite distance from the solid. In a molecular solid, molecules interact only through weak van der Walls forces, so that it is generally accepted that bands are narrow and the highest occupied molecular orbital (HOMO) and lowest unoccupied MO (LUMO) are only slightly displaced from their value in individual molecules.

The number of interest is therefore the distance that separates the HOMO and LUMO levels from the vacuum level, that is, the ionization potential IP and electron affinity EA. Because IP and EA are very sensitive to surface state, their estimation is a difficult task. Several techniques require ultra high vacuum, such as ultraviolet (UV) photoelectron spectroscopy. Other techniques comprise Kelvin probe and electrochemical measurements. In the latter case, care must be taken because several corrections, involving solvatation effects, must be brought to the crude data. Finally, we have to mention internal photoemission, which is in principle a direct method but presents a major drawback: it requires the fabrication of a metal–semiconductor junction and its reproducibility is often uncertain.

In principle, the value of EA can be obtained directly by inverse photoemission spectroscopy, a method that suffers from radiation damage of the organic material. Nevertheless, the most popular way for estimating EA consists of subtracting from IP the HOMO–LUMO gap as deduced from optical absorption measurements (optical gap E_g^{OPT}). However, it must be noted that the optical gap may differ from the true energy gap (corresponding to the energy required to promote an electron from the HOMO to the LUMO level). To understand this difference, we have to go a little deeper in the description of energy levels in a molecular solid. As stated, the IP and EA of the solid slightly differ from that of the isolated molecule. The difference comes from the polarization energy (P_+ and P_- for holes and electrons, respectively) induced by the ionized molecule in the surrounding molecules. This energy is sometimes referred to as the "polaron binding energy." Similar stabilization energy exists for an electron–hole pair (or exciton), which is also often called the exciton binding energy. Only if this exciton binding energy strictly equals $P_+ + P_-$ does the optical gap equal the electronic gap.

When the metal and the organic layer are put in contact, the VL at the interface is shared by both media and barriers for hole (of height Φ_B^p) and electron injection (of height Φ_B^n) appear. For the perfect contact depicted in Figure 2a, the magnitude of the barrier height follows the Schottky–Mott rule, where Φ_m is the work function of the metal:

$$\begin{aligned} \Phi_B^p &= \mathrm{IP} - \Phi_m \\ \Phi_B^n &= \Phi_m - \mathrm{EA} \end{aligned} \tag{1}$$

In molecular solids, which are primarily insulators, the Fermi level is located somewhere in between the HOMO and LUMO levels. With a Fermi level at exactly midgap, we see that the junction in Figure 2a is not in electrical equilibrium. However, equilibration would require charge redistribution, which is difficult in practice because the number of available mobile charges in the organic layer is very low. This is the more true as most OLEDs are made of very thin organic layers, which reduces further the amount of charge available. For these reasons, the simple scheme of Figure 2a remains the one most usually adopted to describe energy level alignment in organic diodes.

Nevertheless, Figure 2a only corresponds to the ideal case. In actual systems, a dipole layer may form at the interface. Possible origins of this dipole are chemical reaction, charge transfer across the surface, or rearrangements of electronic charge. An abrupt shift Δ of the energy levels is associated with this dipole layer, as shown in Figure 2b, which leads to a similar change of the barrier heights. Note that this shift can also be viewed as a change of the metal work function that can be used to purposely adjust the energy level alignment.

From the previous analysis, it emerges that several strategies can be envisioned to tune the energy levels at both sides of a metal–organic junction.

- Interpose a buffer layer between the electrode and the electroluminescent layer to reduce the barrier height. In practice, a hole injecting material presents a high ionization potential, while an electron injecting material is characterized by a low electron affinity. This strategy was proposed in 1987 by Tang and Van Slyke [2] and many of such materials have been identified since then.
- Modify the work function of the electrode by an appropriate surface treatment. Various examples of such treatments will be described in the following.

The energy scheme of a typical multilayer OLED is shown in Figure 3. The terms "hole transport layer" (HTL) and "electron transport electron" (ETL) were first introduced by Adachi et al. [7] in 1988 and are widely used today to designate the two buffer layers. However, we note that these terms are somewhat confusing because these layers *inject* rather than transport the charges from the electrodes to the light-emitting medium. These injecting layers may also play the role of blocking layers for the charges of opposite sign.

Figure 3 does not account for the electric field that appears in the device when applying an external voltage. This is shown in Figure 4 in the case of a one-layer device. The far right image of this figure corresponds to forward bias where emission of light occasionally occurs.

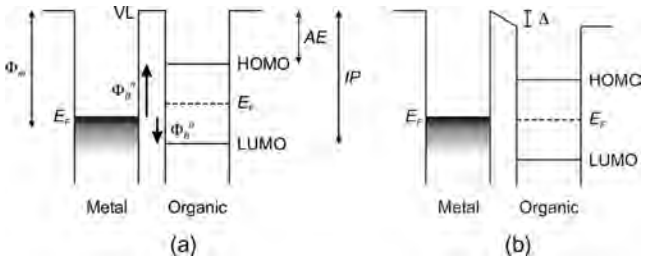


Figure 2. Energy level alignment at a metal/organic interface: (a) ideal case; (b) in the presence of a surface dipole.

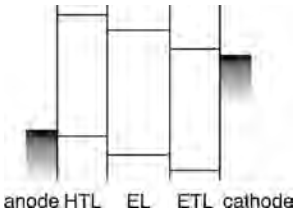


Figure 3. Energy diagram of a multilayer OLED. HTL: hole transport layer; EL: emissive layer; ETL: electron transport layer.

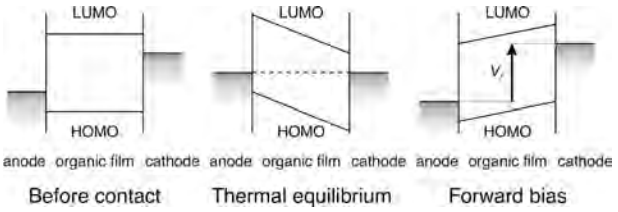


Figure 4. Energy diagram of an unbiased and biased one-layer OLED.

Injection Processes The ideal case of perfect energy alignment at both interfaces is practically never fulfilled in practice. In most real cases, charges have to pass a potential barrier before penetrating the organic layer. This interfacial process must not be confused with charge transport, which will be dealt with in a following section. As most organic semiconductors (SCs) actually behave as insulators, there are basically two injection processes: Schottky emission and field emission. The former corresponds to thermionic emission across the metal–insulator interface. The corresponding current density is given by Eq. (2). Here, A^* is the effective Richardson constant, F is the electric field, q is the absolute electron charge, ε is the permittivity of the organic layer, k is Boltzmann's constant, and T the absolute temperature. β is a characteristic factor given by $\beta = \sqrt{q/4\pi\varepsilon}$, where ε is the permittivity of the medium. (Note that the same constant is encountered in the Poole–Frenkel effect that will be dealt with in the next section.)

$$j = A^* T^2 \exp\left(-\frac{q\Phi_B}{kT} + \frac{q}{kT}\beta\sqrt{F}\right) \tag{2}$$

The second mechanism, which is often referred to as the Fowler–Nordeim process, occurs under high applied electric field. In that case, electrons can tunnel through the triangular barrier that forms when the diode is polarized. The probability for tunneling depends strongly on the height and width of the barrier but is independent of the temperature. The corresponding current density is given as follows, where m^* is the charge effective mass:

$$j \propto F^2 \exp\left(-\frac{4\sqrt{2m^*}(q\Phi_B)^{3/2}}{3e\hbar F}\right) \tag{3}$$

A useful way of checking whether field-induced injection is involved consists of plotting the logarithm of I/F^2 as a function of the reversed electric field.

The main difference between the two mechanisms resides in their temperature dependence; the tunnel process is temperature independent, while the Schottky mechanism is thermally activated. It must be pointed out that, because parameters involved in Eqs. (2) and (3) such as the Richardson constant and effective mass are not well known in most organic SCs, a quantitative check of the charge injection process remains hazardous.

2.1.2. Charge Transport

Although organic SCs have been identified for more than 50 years, charge transport in these materials is still the subject of great debate. A prominent issue of charge transport in organic SCs is the very low mobility of the charge carriers. The main methods for measuring charge carrier mobility are time of flight (ToF), analysis of the current–voltage curves of single carrier diodes in the space charge-limited current (SCLC) regime, and fabrication of a field-effect transistor (FET) structure. As shown in Table 1 for the typical case of PPV, the hole mobility stands in the 10^{-8}–10^{-6} cm^2/V s range. Even lower values have been reported for electron mobility.

This low mobility implies that the corresponding mean free path is substantially shorter than the intermolecular distance, which excludes the dominant charge transport mechanism encountered in most metals and conventional semiconductors, that is, diffusive transport in delocalized states. Among the various models that have been put forward to explain this very low mobility, two classes emerge. One covers the models based on trapping, which assumes that charge transport is limited by a given distribution of traps in the energy gap. Second are the field-dependent mobility models, which implies an exponential dependence of the mobility on the square root of the electric field. The mobility model was first set on the basis of the observed mobility-field dependence of time-of-flight data. As the general dependence resembles that of the Poole–Frenkel mechanism, the model is often referred to as the "Poole–Frenkel-*like*" law because experimental data do not quantitatively agree with the classical Poole–Frenkel emission model.

Because the density of mobile charges in organic SCs is extremely low, the current rapidly becomes limited by charges injected from the electrodes. In this space charge limited regime, the current density is given as follows, where μ is the charge mobility and d is the thickness of the organic layer:

$$j = \frac{9}{8}\mu\varepsilon\frac{V^2}{d^3} \tag{4}$$

The SCLC is strongly affected by the presence of charge traps. The expression depends on the energy distribution of the traps. One of the most frequently invoked distributions is the exponential distribution. If the density of trapped charges largely exceeds that of free charge and the Fermi occupancy function is a step function, the SCLC is given by [11]

$$j = q^{l-1}\mu N_c\left(\frac{2l+1}{l+1}\right)^{l+1}\left(\frac{l}{l+1}\frac{\varepsilon}{H_b}\right)^{l}\frac{V^{l+1}}{d^{2l+1}} \tag{5}$$

Here, H_b is the density of traps and $l = T_c/T$, where T_c is a characteristic temperature that describes the steepness of the distribution.

Table 1. Hole mobility of undoped PPV determined by various methods.

Polymer	Mobility (cm^2/V s)	Method [Ref.]
PPV	10^{-8}	ToF [8]
PPV	10^{-7}–10^{-8}	FET [9]
MEH-PPV[a]	5×10^{-6}	SCLC [10]

[a] MEH-PPV—poly[2-methoxy,5-(2′-ethylhexyloxy)-1,4-phenylene vinylene].

The field-dependent mobility model dates back to the 1970s [12]. It has become common practice to analyze mobility data in terms of the empirical equation due to Gill [13]:

$$\mu(F) = \mu(0)\exp\left(-\frac{\Delta_0}{kT} + \frac{q}{kT_{\text{eff}}}\beta\sqrt{F}\right) \qquad \frac{1}{T_{\text{eff}}} = \frac{1}{T} - \frac{1}{T_0} \tag{6}$$

T_0 is the temperature at which Arrhenius plots of $\mu(T)$ intersect and β is the previously defined Poole–Frenkel factor. The physics of the field dependence relies on disorder-induced localization of charges. The field-dependent term arises from the fact that charge transport is governed by tunneling from one localized state to another; the transfer occurs through tunneling, a process that can be assisted by the electric field. Such a mechanism is similar to the Poole–Frenkel effect. However, the exact expression for β strongly depends on the distribution of the traps, so that it noticeably differs from the conventional Poole–Frenkel factor. Furthermore, the derivation of the constant cannot be carried out by analysis. A widely used expression is that derived by Bässler [14] by performing a statistical Monte Carlo computation based on a Gaussian distribution of localized levels. The *energetic* disorder is depicted by the standard deviation σ, while that of the *geometric* dispersion is given by Σ. The general expression of the field-dependent mobility is given by

$$\mu = \mu_0 \exp\left[-\left(\frac{2}{3}\frac{\sigma}{kT}\right)^2\right] \exp\left\{C\left[\left(\frac{\sigma}{kT}\right)^2 - \Sigma^2\right]\sqrt{F}\right\} \tag{7}$$

More recently, Garstein and Conwell [15] and Novikov et al. [16] have proposed an alternative approach to the problem. The model, termed the correlated disorder model, assumes that the potential viewed by a charge carrier presents slowly varying spatial fluctuations. These variations occur because the energy of neighboring sites is correlated; that is, the energy of one site depends on the energy of the nearest sites. Because of these correlations, the energy difference between two sites is limited by an effect of spatial averaging. As in the case of the model of Gaussian disorder presented previously, the mobility is calculated through numerical Monte Carlo simulation:

$$\mu = \mu_0 \exp\left[-\left(\frac{3}{5}\frac{\sigma}{kT}\right)^2\right] \exp\left\{C_0\left[\left(\frac{\sigma}{kT}\right)^{3/2} - \Gamma\right]\sqrt{\frac{aF}{\sigma}}\right\} \tag{8}$$

Here, $C_0 = 0.78$ in a three-dimensional case, Γ measures the energy disorder ($\Gamma = 2$), and a is the intersite distance.

2.1.3. Charge Recombination and Light Emission

Charge Recombination The problem of charge recombination has received much less attention than charge injection and light emission. Charge recombination is a bimolecular reaction that obeys a second order kinetic of the following form, where n and p are the density of electron and holes, respectively [17]:

$$\frac{dn}{dt} = \frac{dp}{dt} = \gamma np \tag{9}$$

The most widely recognized mechanism of charge recombination is of Lagevin type, that is, diffusion controlled [18]. Such a mechanism is characteristic for materials in which the charge carrier mean free path is smaller than a critical distance defined by the Coulombic capture radius $r_c = q^2/4\pi\varepsilon kT$. For typical organic emitting materials, r_c amounts to around 20 nm at room temperature, while the mean free path compares to the distance between conjugated sites, which is of the order of a few nanometers. The recombination factor in Eq. (9) is given as follows, where μ_n and μ_p are the electron and hole mobility:

$$\gamma = \frac{q}{\varepsilon}(\mu_n + \mu_p) \tag{10}$$

From Eqs. (9) and (10) it can be seen that increasing recombination can be obtained by increasing charge mobility but also requires the density of electrons and holes be as close to each other as possible. The latter issue demands a balanced injection rate for both charges, which is probably one of the more difficult requirements for high performance devices.

Light Emission Charge recombination leads to the formation of electron–hole pairs that rapidly decay to excitons. The nature of the excitons in conjugated organic materials is the subject of great controversy that is still alive today. The debate has been the subject of a whole volume [19] published in 1997, which develops the two opposing aspects of the controversy. In essence, an exciton is an electron–hole pair where the two charges are linked together through Coulomb interaction. The strength of the link, also called exciton binding energy, is strongly material dependent. In conventional inorganic semiconductors the effective mass is lower than the free electron mass and the dielectric constant is high, so the binding energy is weak (that is, comparable to kT at room temperature) and the exciton extends over several lattice sites. (This weakly bond exciton is frequently termed Wannier exciton.) Hence, exciton effects are unimportant except at low temperature. On the other hand, excitons in molecular crystals are tightly bound and are localized on a single molecule, thus forming a so-called Frenkel exciton [20]. Because of the weakness of intermolecular interactions, the Frenkel exciton model is widely accepted in materials made of small molecules. This is not the case for polymers, where there exists a conflict between strong *intra*chain and weak *inter*chain interactions. Estimating the exciton binding energy is in principle quite simple [21]; it merely amounts to the difference between the energy gap and the optical gap. However, there are several difficulties hidden behind this simple definition. The first one is that both quantities are chain length dependent. Only for sufficiently long chains do they come to limiting values independent of the length. Another problem arises from the tendency of conjugated chains to relax when charges are added. This point has already been evoked earlier in this chapter. A consequence of this relaxation is that the energy gap is reduced by a quantity $P_+ + P_-$, where P_+ and P_- are

polarization energies, also called polaron binding energies. Because the polarization energies amount to tenths of an electronvolt, the estimation of the exciton binding energy may substantially differ depending on whether, on breakup, the exciton dissociates directly into a pair of polarons rather than a free electron and a free hole that subsequently relaxes into polarons. Part of the controversy rests on this still unresolved point.

Frenkel excitons in organic semiconductors can also be viewed as excited molecules that can travel over the solid through a mechanism of energy transfer. Because the Frenkel exciton propagates, it can generate energy bands, the width of which is determined by the transfer rate from one molecule to the next.

Frenkel excitons can be singlet or triplet, depending on the respective spin parity of its constituents. During the mechanism of electroluminescence, excitons are a result of the recombination of two charges of reverse sign, the spin of which is distributed at random. A generally accepted concept is that this feature limits the electroluminescence yield to 25%. The argument is based on simple spin multiplicity; for recombination of a pair of charges, they are four microstates in total, with three triplet states and only one singlet state. As the ground state is usually singlet, radiative decay from a triplet state is spin forbidden, so that the theoretical quantum yield of OLEDs is limited to 25%. However, quantum efficiency up to 50% has been recently reported in MEH-PPV based diodes [22]. Two kinds of argument have been invoked to explain this behavior.

The first one is to assert that excitons in some organic semiconductors (mainly polymers) are weakly bound (Wannier excitons). In that case, exciton effects can be neglected and quantum yields up to 100% be envisioned, like in the case of inorganic LEDs. Such an argumentation falls within the general controversy evoked previously.

The second analysis assumes Frenkel-type excitons [23]. The ratio between the number of singlet excitons and the total number of excitons can be expressed as $\eta = \sigma_S/(\sigma_S + 3\sigma_T)$, where σ_S and σ_T are the cross sections for singlet and triplet formation. The quantum mechanical argument is that interchain correlation effects of bond-charge type distinguish the cross sections for singlet a triplet states. Calculations show that in the case of PPV, the ratio σ_S/σ_T may be as high as 3, thus leading to a quantum yield of 50%.

An alternative approach to improve quantum efficiency consists of using phosphorescent materials [24]. In these materials, the luminescence comes from the radiative decay of the triplet state. In that case, the theoretical limit of the quantum yield is 75%. One way would therefore consist of adding phosphorescent compounds to the luminescent layer of the diode. However, this method has many problems. Because the decay from a triplet state is forbidden by selection rules, the lifetime of triplet excitons is very long. This means that all the phosphorescent sites may become occupied and therefore unavailable for energy transfer. Moreover, the long lifetime also increases the risk of nonradiative deactivation at some defect, or through triplet–triplet interactions.

A group at Princeton University [25] has used a different methodology where both singlet and triplet excitons transfer energy from the electron–hole pairs. As shown in Figure 5,

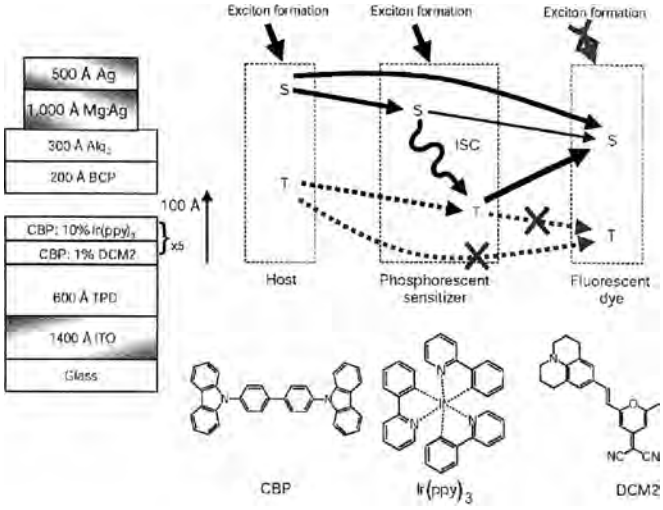


Figure 5. Phosphorescent sensitized OLED. Left: schematic structure of the device. The structure comprises a HTL (TPD), a hole blocking layer (BCP) and an ETL (Alq_3). The emitting layer is made of 10 layers of CBP alternatively doped with $Ir(ppy)_3$ (10%) and DCM2 (1%). Right: proposed transfer mechanism in the emitting layer. S stands for singlet, T for triplet, and ISC for intersystem crossing. Förster transfers are represented by solid arrows and Dexter transfers by dotted arrows. Reprinted with permission from [25], M. A. Baldo et al., *Nature* 403, 750 (2000). © 2000, Macmillan Magazines Ltd.

the emissive layer is made of a set of alternating layers that contain either a phosphorescent sensitizer or a fluorescent dye embedded in a host material. As a result, both singlet and triplet excitons are generated. Moreover, the energy of the triplet state in the phosphorescent dye is higher than that of the singlet in the fluorescent dye, so that energy transfer occurs from the triplet to the singlet, which then emits light. A quantum yield of nearly 100% thus has be obtained. However, the system presents several drawbacks. One is the need for a number of downhill energy steps, which results in the energy of the emitted light being much lower than that of the charge pair. For this reason, green or blue light would be difficult to obtain. The second problem concerns the energy transfer. Energy transfer between molecules in a solid may occur through two basic mechanisms: dipolar (Förster) or contact (Dexter). The former is long range and authorizes change of spin configuration, while the latter is short range and an exciton transferred this way retains its spin multiplicity. This second mechanism should therefore be avoided because it would lead to the formation of the nonemitting triplet exciton in the fluorescent dye.

2.2. Materials

One of the unique aspects of organic chemistry is its ability to produce new molecules at a high rate. Organic electroluminescence is a very active research topic that involves a great number of teams in both academic and industrial laboratories worldwide. Since the emergence of organic electroluminescence, dozens of new compounds aimed at being used in OLEDs have been synthesized each year. For these reasons, a comprehensive review of the materials used in organic LEDs would be far beyond the scope of this chapter. It would also not be very helpful, because it is difficult

to predict which compounds will lead to real applications. We will therefore restrict ourselves to a few representative molecules. Note that many of them are already commercially available.

2.2.1. Emitting Materials

Emitting materials can be sorted into polymers and short molecules. Recall that organic electroluminescence was first observed in a molecular crystal, anthracene, in 1963. This early device already presented encouraging quantum efficiency. However, its operating voltage of around 1000 V made its development unlikely. It is only with the advent of thin film devices that commercial application could be envisioned. Polymers [26] and small molecules [27] based devices emerged almost simultaneously at the turn of the 1980s. This breakthrough was actually made possible after the synthesis of the first conducting polymer in 1977: polyacetylene (PA) [28]. At that time, the unique property that fascinated chemists and physicists was the ability of the polymer to become conducting upon doping by a charge donating of withdrawing agent, hence the name "conducting" polymers. PA has a serious drawback: it is highly unstable in ambient air, but many other conjugated polymers were soon discovered. Their molecular formula is given in Figure 6. The finding of electroluminescence in one of them, PPV, drastically changed the perspective. The more promising property of this class of materials moved from "being a conductor when doped" to "being a semiconductor when undoped." A very interesting issue in conducting polymers is the possibility of substituting side groups to the main chain in order to modify some of its physical properties. For example, unsubstituted PPV is infusible and insoluble in most organic solvents, thus making it very difficult to process in thin solid films. Some of its substituted compounds are sketched in Figure 7. Now the most popular, MEH-PPV, which is soluble, has largely helped in developing OLEDs based on PPV. Similarly, it has been theoretically predicted that substitutions with electron donating or withdrawing groups could allow a tuning of the frontier (HOMO and LUMO) levels [29], thus favoring energy level alignment at one or both electrodes. The HOMO and LUMO shifts calculated for compounds **I** and **II** in Figure 7 are displayed in Table 2.

Substitution with a donor group (OCH_3, compound **I**) leads to a destabilization (positive shift) of the levels, while substitution with an acceptor group (CN, compound **II**) induces stabilization (negative shift).

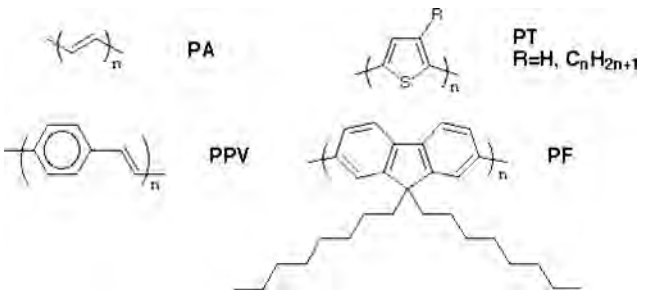


Figure 6. Molecular structure of most popular conducting polymers. *Trans*-PA; polythiophenes; PPV; poly-dialkyl-fluorene.

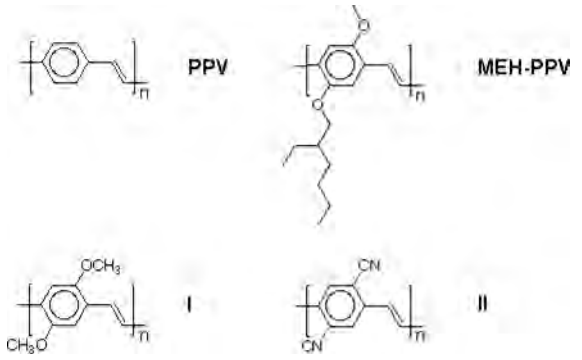


Figure 7. Molecular structure of PPV and poly[2-methoxy,5-(2′-ethylhexoxy)-1,4-henylene-vinylene] (MEH-PPV). Compounds **I** and **II** are models used to see the effect of electron donor and acceptor on the frontier levels (see Table 2).

A critical issue with low molecular weight organic compounds stems from their tendency to crystallize readily and hence form polycrystalline layers when deposited in thin films. This has the following detrimental effects on the performance of EL diodes: (1) exciton recombination occurs at grain boundaries, which considerably reduces the quantum yield; (2) films are not free of pinholes, leading to short circuits; (3) although data on that matter are relatively scarce, the polycrystalline structure might also be responsible for the lack of long-term stability of the devices. The original structure developed by Tang and van Slyke was made of two layers. The first one was an aromatic diamine and the second one was tris(8-hydroxyquinoline) aluminum (Alq_3). The morphology of the films, as measured by transmission electron microscopy, was found to be amorphous for the first layer and microcrystalline for the second one. This probably explains the tremendous success of Alq_3 as an emissive layer in OLEDs made of small molecules. The molecular structure of Alq_3 is given in Figure 8. More recently, Shirota [30] developed a general strategy to conceive and synthesize amorphous molecular materials. Starting from the simple picture that nonplanar structures should prevent crystal packing, Shirota proposed a new concept, the π-electron starburst, for the design of amorphous molecular materials. Typical examples are shown in Figure 9. (Note that Alq_3 also presents a nonplanar configuration.)

2.2.2. Electrodes

The basic rules for choosing materials for electrodes are quite simple: high work function for the anode (hole injecting) and low work function for the cathode (electron injecting). To that, we must add an obvious requirement; that is,

Table 2. Calculated shifts (in eV) of the HOMO and LUMO levels for derivatives **I** and **II** with respect to the unsubstituted PPV.

Derivative	HOMO	LUMO	E_g
PPV			3.24
I	+0.27	+0.08	3.06
II	−0.99	−1.17	3.11

Note: The energy gap E_g reported in the last column corresponds to the difference between the HOMO and LUMO levels (after [29]).

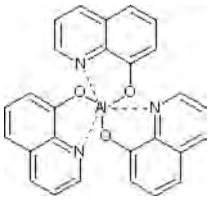


Figure 8. Molecular structure of tris(8-hydroxyquinoline aluminum).

at least one of the electrodes must be transparent to let light get out of the device. Transparent conducting materials mainly include degenerate oxides. Although there have been some reports on OLEDs with aluminum-doped zinc oxide (AZO) [31], practically all OLEDs use indium-tin oxide (ITO) as the anode. A probable reason for that is that ITO coated glass is also widely used in liquid crystal displays. In most cases, the organic layers are deposited on top of ITO during the fabrication of the device. In the course of the process, the surface state of the electrode is of primary importance because it controls crucial parameters such as energy level alignment and the absence of short circuits. In a recent paper, Kim and co-workers have investigated how surface treatments modify the performance of OLEDs [32]. The various treatments and their effect on the work function and roughness are summarized in Table 3.

The desired issues are (1) high work function to improve hole injection in the organic layers; (2) low sheet resistance to reduce ohmic losses; (3) low average roughness to avoid short circuits. An analysis of the data in Table 3 shows that the oxygen plasma treatment appears to be the best one. This was further confirmed by measuring the lifetime of the diode, which was substantially improved after an oxygen plasma treatment of the ITO surface.

Conducting polymers such as polyaniline [33] have been envisioned as alternative transparent anodes. The polymers present the advantage of a better interfacial matching with organic semiconductors. However, the sheet resistivity of these polymers remains too high for practical use. Hybrid anodes that combine the higher conductivity of ITO with the properties of the polymer have been shown to bring improvements in terms of efficiency, lifetime, and turn on voltage. The best example is the ITO/poly(3,4-ethylene dioxythiophene) (PEDOT) anode. PEDOT is now a commercially available polymer. In its most widely used form,

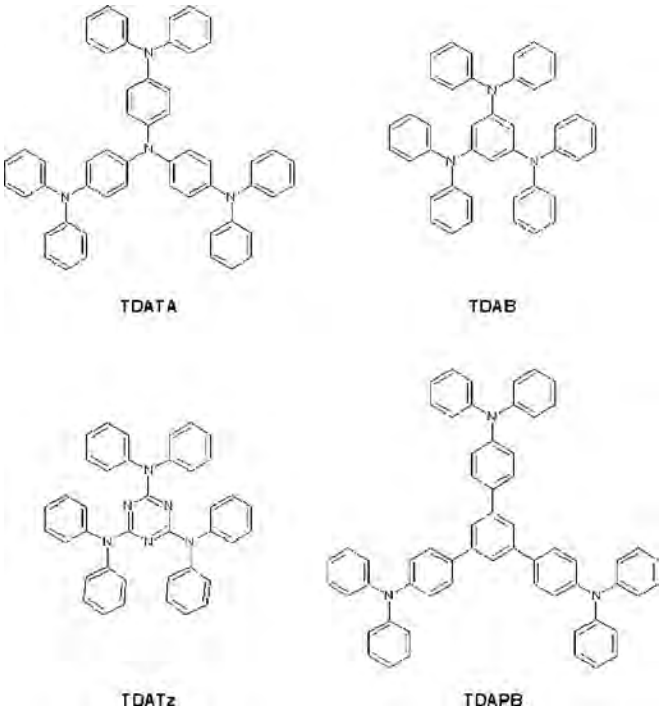


Figure 9. Molecular structure of typical amorphous molecular materials: 4;4′,4″-tris(diphenylamino)triphenylamine (TDATA); 1,3,5-tris(diphenylamino)benzene (TDAB); 2,4,6-tris(diphenylamino)-1,3,5-triazine (TDATz); 1,3,5-tris(4-diphenylaminophenyl)benzene (TDAPB).

the polymer is doped with another polymer, polystyrenesulfonate (PSS), Figure 10. A unique feature of this mixture is its large stability under its doped conducting form. Electroabsorption measurements [34] have shown that the improvements partly originate from an increased work function, so that PEDOT also plays the role of a hole injecting material.

An alternative way of modifying the work function of ITO is the use of self-assembled monolayers (SAMs). The principle of this modification rests on the use of dipolar molecules. Molecules used in SAMs also comprise a functional group prone to induce chemisorption on the surface. In the case of ITO, it has been shown that both carboxylic acid and preferably phosphonic acid can induce the formation of SAMs [35]. Both types of SAMs have been used to modify the ITO

Table 3. Work function, sheet resistance, and roughness of variously treated ITO substrates.

	Surface treatment	Work function (eV)	Sheet resistance (Ω/□)	Roughness (nm)
	as-received	4.5	16.1	2.6
Mechanical	paper rubbing	4.2	16.3	2.3
	Teflon rubbing	4.2	16.5	2.4
	ultrasonic	4.35	15.5	3.4
Wet	RCA	4.35	19.6	2.4
	aquaregia (10′/20′/30′)	4.6/4.3/4.7	18.5/23.5/28.6	3.8/8.4/8.8
Dry	oxygen plasma (5′/10′/15′)	4.35/4.75/4.75	16.4/15.0/16.4	1.4/1.4/2.1
	argon plasma (5′/10′/15′)	4.5/4.5/4.55	16.7/17.3/17.0	10.9/15.4/23
Combined	aquaregia (20′)/oxygen plasma (10′)	4.6	27.7	6.0
	oxygen plasma (10′)/aquaregia (20′)	4.7	>30	1.8

Note: After [32].

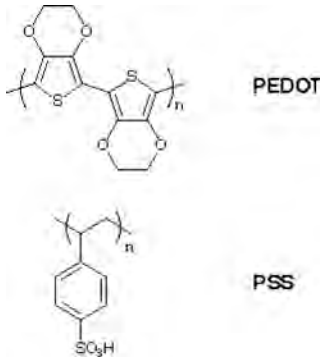


Figure 10. Molecular structure of PEDOT and PSS.

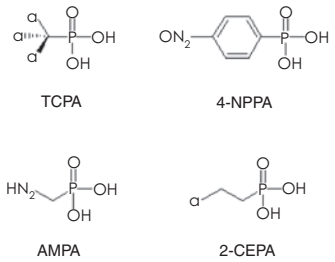


Figure 11. Molecular structures of the SAMs used to modify ITO electrodes: (tirchloromethyl)phosphonic acid (TCPA); (4-nitrophyl)phosphonic acid (4-NPPA); (aminomethyl) phosphonic acid (AMPA); (2-chloroethyl) phosphonic acid (2-CEPA).

electrodes in OLEDs [36, 37]. The shift in work function is interpreted as originating from the dipole moment induced by the SAM. According to this model, the work function shift $\Delta\varphi$ is determined by the change in electrostatic potential ΔV created at the surface by the dipolar moment μ,

$$\Delta\varphi = -q\Delta V = -qN\frac{\mu}{\varepsilon} \tag{11}$$

Here, N is the surface density of molecules and ε the permittivity of the SAM. Table 4 compares the calculated and measured work function change by four various molecules adsorbed on ITO. The measurements were performed with a Kelvin probe. The molecular structures of the SAMs are given in Figure 11.

Because the cathode is most generally deposited at the end of the fabrication process, surface treatments cannot be performed and the only criterion for selecting a metal for the cathode is its work function, which must be as low as possible (electronegative metal). Table 5 gives the work function of various metals used as cathode. According to Table 5, metals such as calcium or sodium appear to be the best choice. Unfortunately, these metals are also highly reactive and making devices with them requires a very efficient protection against ambient air. To date, the best compromise seems the use of an alloy between low work function and noble metal, such as Mg:Ag.

Improved electron injection was obtained by inserting a thin (0.5–10 nm) insulating film of lithium fluoride (LiF) between a stable metal cathode (mainly Al) and the emitter layer. The mechanism behind the enhanced injection is not fully understood. Interpretations include tunnelling [38], removal of interface states at the cathode–organic interface, and reduction of the barrier height to electron injection [39].

2.2.3. Charge Transport Agents

Although the terms "electron transport" and "hole transport" layers are universally used in the community of organic light-emitting diodes, it must be stressed that these expressions are somewhat misleading. In essence, an electron (hole) transport material would be that in which high electron (hole) mobility is encountered. This is not the case in OLEDs. Instead, ETLs and HTLs are made of chemical agents that favor electron (hole) injection. For this reason, some recent papers make a contradistinction between charge *transport* and charge *injecting* materials. In reality, both terms designate the same class of materials, since charge carrier mobility is low in practically all these organic semiconductors.

One of the most widely used electron transport agent remains the already invoked Alq_3, which is most often used as both ETL and emissive layer. A few alternatives have been proposed since then. At present, the main representatives of electron transport agents are derivatives of oxadiazole. One of these, PBD, is shown in Figure 12. The difficulty in finding organic electron transport materials probably comes from the fact that low electron affinity is usually associated with high chemical reactivity. Note that some polymers, such as polypyridine [40], have also been reported as electron transporting agents.

Hole transport materials are much more common. These comprise both short molecules and polymers. A very popular hole transport agent is N,N'-diphenyl-N,N''-bis(3-methylphenyl)(1,1′-biphenyl)-4,4′-diamine or TPD [7], the molecular formula of which is indicated in Figure 13.

Table 4. Calculated and measured work function shift (in eV) of ITO modified by SAMs.

SAM	Dipole moment (D)	Calculated	Measured
4-NPPA	5.73	0.303	0.720
TCPA	1.76	0.185	0.221
2-CEPA	1.69	0.179	0.212
AMPA	−1.43	−0.140	−0.179

Note: The molecular structure of the SAMs is shown in Figure 11. Dipole moments are calculated at the PM3 semi-empirical method (taken from [37]).

Table 5. Typical low work function metals.

Metal	Work function (eV)
Al	4.28
Ca	2.87
In	4.12
Li	2.9
Mg	3.66
Na	2.75
Sm	2.7

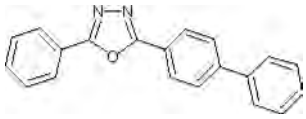


Figure 12. Molecular scheme of 2-(4-biphenylyl)-5-phenyl-1,3,4-oxadiazole (PBD).

2.3. Performance

To conclude this section, we give here a brief outlook of the current performance of organic light emitting diodes.

2.3.1. Color

It is often claimed that one of the unique advantages of OLEDs is the possibility of generating practically all colors. However, this statement must be somewhat tempered. A characteristic of organic luminophores is to present a wide emission spectrum. This is detrimental when dealing with color displays, where red, green, and blue emitters with well-defined chromaticity are needed, but turns out to an advantage in the case of white light emission. These two extreme cases will now be dealt with in sequence.

The development of color displays based on inorganic light-emitting diodes has long been impeded by the lack of blue emitting devices. By contrast, blue OLEDs have been realized since the early beginning [41]. To date, dozens of blue light emitting molecules and polymers have been identified. Strangely enough, the problem with OLEDs concerns the red emitters; most of them are actually orange, which is not favorable for making good quality color displays. Two approaches have been used to overcome this difficulty. One is to use organometallic complexes based on rare earths, mainly europium, which is one of the best-known red emitters [42, 43] with four sharp emission peaks, the dominant one being centered at 618 nm. It is worth pointing out that in this case, the emitting element is the rare earth ion, which is at variance with other organometallic complexes such as Alq_3 where the emitting part is the ligand. The sharp emission in rare earths comes from electron transitions between $4f$ bands. For symmetry reasons, these transitions are forbidden in the free ions but may become allowed when the symmetry of the ion is removed by an asymmetrical external crystal field such as that provided by an appropriate ligand. Generally, the complex is embedded in a polymer matrix that can act as an energy transferring system for the complex. After exciting the polymer, energy cascades through the ligand to the central rare earth ion, which eventually emits light. We note that organic diodes have also be made with another rare earth, erbium [44, 45], that emits in the near infrared at 1.54 μm, a wavelength of importance in telecommunications because it corresponds to a minimum of losses in optical fibers.

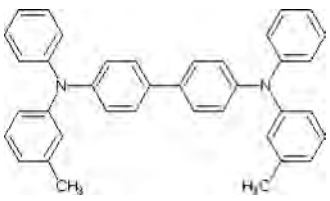


Figure 13. Molecular formula of N,N'-bis(3-methylphenyl)-N,N'-diphenylbenzidine.

The second strategy for red light devices is the use of small amounts of an efficient luminophore for this color dispersed in an electroluminescent material that emits at higher energy. One speaks in that case of "dopant" (or guest) in a host material. Representative examples of doping materials are laser dyes such as 4-dicyanomethylene-2-methyl-6-(*p*-dimehylaminostyryl)-4H-pyran (DCM) [46]. Like in the case of the rare earth complexes, emission results from a Förster energy transfer from the matrix (typically Alq_3) to the dye.

The first report on white OLED was published in 1994 [47]. White light is obtained by mixing several (most often three) emitting materials with complementary emitting spectra. In this configuration, the wide luminescent spectra of organic compounds become advantageous. Various structures have been used to realize white OLEDs, among them blends, multiple layers, and doped matrices. In essence, the basic principles that govern the operating mode of these diodes do not differ from those of the one-color diodes. Note that white phosphorescent diodes have also been reported [48].

2.3.2. Yield

The yield of an electroluminescent device can be defined in various ways. The quantum yield η_q represents the ratio of the number of emitted photons to the number of injected charge carriers. The overall yield can be decomposed into a product of factors that mirror the yield of each step of the process; hence $\eta_q = \eta_1\eta_2\eta_3$, where η_1 = (number of emitted photons)/(number of singlet excitons), η_2 = (number of singlet excitons)/(total number of excitons), and η_3 = (number of excitons)/(number of injected carriers). The first term corresponds to the photoluminescence (PL) quantum efficiency of the emitting material. Note that high PL yield (higher than 50%) can be achieved in organic solids. As discussed, the second term is in principle limited to 25%, but there is both experimental and theoretical evidence for possible higher yields. Finally, the third term mainly depends on the rate of carrier recombination. A crucial issue for high recombination rate is a balanced injection of electrons and holes.

A distinction has been made by Greenham et al. between the *internal* and *external* yield [49]. To understand this concept, one must recall that all photons produced in the diode do not actually leave the emitting layer; some of them are trapped by a process of total internal reflection. Using Snell's law, they determined that the internal quantum yield, which accounts for the total number of photons produced, is greater than the external quantum yield, for directly emitted light, by a factor of $2n^2$ where n is the refractive index of the emitting material. Note that this correction is much more important in inorganic semiconductors than in the organic ones, the refractive index of which generally ranges between 1.5 and 2.

However, the technologically important parameter is the power efficiency, that is, the ratio of the output light power (in lumens) to the input electrical power (in watts). It is worth noting that the early device of Pope and co-workers made of an anthracene single crystal presented a reasonably good quantum yield of a few percent but a desperately

poor power efficiency because its operation required very high voltages. Current OLEDs operate at low voltage, which considerably increases the power yield of the diodes. Among other issues, high power yields are obtained by using very thin films; the overall thickness of current diodes does not exceed 100 nm.

2.3.3. Degradation

Device stability is a major concern for OLEDs. This problem has known increased interest, as the development of organic electroluminescence is becoming an industrial challenge. Studies have focused on two widely used structures: ITO/poly(para-phenylenevinylene)/Ca and ITO/Alq_3/(Al or Mg:Ag) (with occasionally TPD as hole transport layer). Two different degradation mechanisms have been identified [50]. The first corresponds to a uniform long-term decay in electroluminescence yield. The half-time of the best current OLEDs, defined as the time elapsed before the luminance of the device falls to half its initial value, may amount to up to a few thousands of hours. In the second process, the degradation manifests itself through the formation of dark (i.e., nonemissive) spots that gradually cover the surface of the diode [51]. The formation of these dark spots is greatly reduced when the device operates in an inert atmosphere. For this reason, great care must be taken to avoid exposure of OLEDs to oxygen or moisture during their fabrication. Among the various origins for degradation that have been put forward, we can mention:

- Crystallization of the charge transporting layers [52] may occur. As stated, the best performance is obtained with amorphous materials. A likely origin for crystallization is Joule heating. Note that an improvement of the luminescence yield also results in a reduction of the Joule current.
- Evidence has been given that the dark spot formation is associated with the cathode/Alq_3 interface [53]. A major cause is cathode delamination. Another origin for dark spots is oxidation of the cathode/organic interfacial region. The role of traces of oxygen in this process is strongly suspected.
- Degradations may also occur at the ITO/organic interface. The main cause is diffusion of indium out of the oxide electrode. It has been suggested to replace the ITO anode with AZO [31], another transparent conducting oxide. Diffusion of zinc also occurs in this case, but to a lesser extent than indium out of ITO.

A general conclusion of these degradation studies is that preventing exposure of the device to oxygen and moisture during its entire fabrication process is of primary importance. It is also worth pointing out that any improvement in the performance of a device most generally comes with an increase of its lifetime.

3. NANO-ORGANIC LIGHT-EMITTING DIODES

The term "nano-OLEDs" designates organic light emitting diodes having nanometric size (i.e., from several nanometers to several tens of nanometers). Making such devices with controlled properties is still difficult because the fabrication technique of OLEDs has not yet been entirely mastered. Moreover, the technology for nanosize classical semiconductors is still in the process of development, despite remarkable progress made in this field in recent years. Several researchers, however, have tentatively fabricated OLEDs of submicrometer size with moderate success. The basic techniques are borrowed from semiconductor technology, although the particular properties of organic materials could allow special preparation and applications of devices. In 2001, the limiting size for nanodevices was reported to be about 250 nm [54] although it recently has been demonstrated that nanocontacts of 10 nm could be realized by nanoimprint technology [55].

We shall first review the different techniques used in conventional semiconductor technology to fabricate nanosize devices.

3.1. Technology of Nanostructures

3.1.1. Lithography

The key drivers for semiconductor industry have been lithography. This technique employs optical projection printing: the image of the mask is projected on a semiconductor wafer that has been coated with a photosensitive layer or photoresist (Fig. 14). After exposure to UV light, this layer, which becomes soluble in a developing fluid, is removed so that a pattern emerges on the substrate upon development by appropriate solvents. This pattern is a resist layer for further etching or ion implantation process and allows defined geometries for the devices to be obtained. The resolution of the pattern depends on the quality of both the resist material and the mask fabrication. The former should exhibit high imaging contrast whereas the latter should be well defined, resolved, and placed with accuracy.

The resolution of a lithography system is usually expressed as a function of the wavelength λ of the radiation used to

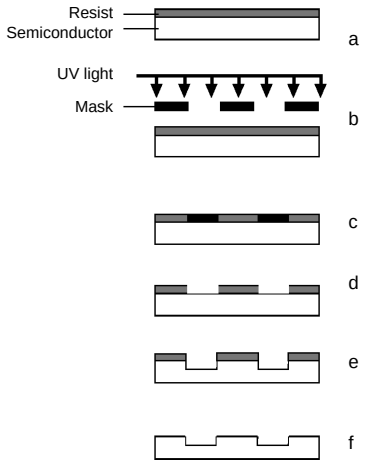


Figure 14. Schematic overview of an optical lithography process: (a) semiconductor covered with light sensitive resist; (b) exposure to UV light; (c) the expose region becomes soluble; (d) reproduction of the pattern on the resist layer; (e) etching of the semiconductor; (f) washing of the resist; the pattern is transferred to the substrate.

project the image and the numerical aperture F,

$$d = k_1 \frac{\lambda}{F} \tag{12}$$

where k_1 is a constant whose value depends on the process being used, typically 0.5 for integrated circuit manufacturing, and F varies in the range of 0.5–0.6. To achieve a low resolution, short wavelengths are used, for instance, with excimer laser sources or very short wavelengths from ultraviolet radiation (100–140 nm). The main drawback of these techniques is a large absorption of the radiation used by the materials, and the resist layer should be relatively thick.

Alternative technique consists of using an electron beam as a radiation source and a membrane with holes in it as a mask (electron beam lithography). The electron beam is absorbed by the solid part of the membrane and passes through the holes in the membrane. In this way, patterns can be designed on the resist layer. The resolution limit is based on the intermolecular force between exposed and unexposed molecules, which prevents the exposed molecules from being dissolved in the developer solution. Heating of the mask by the electron beam may cause distortion in the final pattern, which can be defined with a resolution lower than 100 nm.

3.1.2. New Technologies

The conventional lithography techniques provide highly resolved geometries to devices but require specific preparation steps, which are not always convenient for the fabrication processes. Several techniques recently developed can be used in replacement of traditional lithographic processes with success, especially for organic and polymer based devices.

(a) *Scanning probe arrays:* This technique makes use of the scanning probe of a scanning force microscopy system to pattern the surface of a layer by affecting its surface chemistry. A bias current is applied to the tip of the probe and generates an electrical field in the surface under test and modifies its structure. This modification produces relief images that can be used for circuit pattern. A large number of scanning tips in array arrangement would be needed to design circuits by this technique.

(b) *Nanoimprinting lithography* (*NIL*): This technique uses a stamping or moulding process to replicate a pattern from a master. It consists of transferring a pattern by a mechanical mold to a layer deposited on a hard substrate (Fig. 15). A stamp with the desired nanofeatures is usually fabricated by a conventional lithography technique. The material to be printed is a polymer, which is deposited on a substrate. The stamp and the substrate are put into conformal contact and the system is molded by heating it at a temperature above the glass temperature T_g of the polymer. The stamp and the substrate are then cooled down to a temperature just below T_g and separated. The polymer film left on the substrate can be used as a device or a mask for further patterning operation. Several commercial products are fabricated by using this process, for instance, compact disks. Resolution as small as 10 nm over large areas could be realized by this technique [56]. For practical use, there are two considerations. The first one is the alignment accuracy of the patterns, which can be important especially for small size and highly resolved circuits. The second issue is the accuracy in the size of the mold that should be as good as that of the pattern.

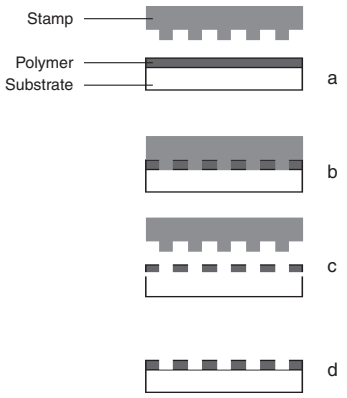


Figure 15. Schematic overview of the nanoimprint technique: (a) initial setup; (b) conformal contact; (c) cooling and mold removal; (d) patterned polymer layer.

3.2. Polymers in Nanotechnology

Polymers are perfectly adapted for nanotechnology because of the length scale of their chains. In addition, ease of processing and the ability to incorporate chemical functionality to the backbones make polymers ideal building blocks for nanosize systems [57].

In the NIL technique, a polymer such as poly(methyl methacrylate) or PMMA is used as a thin thermoplastic film to pattern nanosized designs on semiconductor substrates. After imprinting the polymer film, the pattern is transferred to the substrate by etching. The quality of the process depends on several parameters such as T_g temperature, viscosity of the polymers, and adhesion of the polymer to the substrate [58].

An alternative technique using polymer to design patterns is microcontact printing (μCP). Because the process does not employ hard polymers, it is called a soft lithography technique. It consists of printing a monolayer of a material from a stamp [made of poly(dimethyl siloxane) or PDMS] on a substrate (Fig. 16). The background or left space on the substrate can be later filled with a self-assembled monolayer, producing a patterned surface. The resolution obtained by this technique can reach 100 nm [59].

Several other particular techniques have been developed to pattern electroluminescent devices. They will be described in the next section. We mention here a special technique using an electrical field to pattern thin polymer film [60]. It consists of depositing a thin film on an electrode and then applying an electrical field to the polymer through another electrode, kept at a distance of 100 nm from the surface at the film, the system being kept at a temperature above the T_g temperature of the polymer. The electrical field will induce displacement charge density at the surface of the polymer, producing a periodic pattern.

3.3. Dispersed Nano-OLEDs

Making nanosized LEDs can be carried out by using a polymer host matrix in which electroluminescent nanoparticles are incorporated. These devices are known as hybrid structures and the nanoparticles are responsible for the light emission. The nanoparticles may be organic or inorganic. Guha et al. [61], for instance, used fluorescent dyes such as 3-(2′-benzothiozolyl)-7-*N*,*N*-diethylaminocoumarin (coumarin 7), 3-(2′-benzothiozolyl)-7-diethyl-aminocoumarin (coumarin 6), and DCM dispersed in a PMMA matrix. The thin films obtained by conventional deposition techniques were brought into contact with a GaN based LED. Light emission from the inorganic material is absorbed by the organic particles, which in turn emits light. Alternatively, inorganic nanocrystals can also be used in a electroluminescent polymer to obtain light emission from both materials. Schlamp et al. [62] deposited cadmium selenide (CdSe) or cadmium sulphide (CdS) nanocrystals from a chloroform solution onto a PPV thin film to form a disordered matrix of nanocrystals. Devices made with the combination of the two layers show either nanocrystal-only EL or a combination of nanocrystals and polymer EL, depending on the nanocrystal layer thickness. Several other inorganic nanoparticles have also been used in a polymer matrix in order to fabricate electroluminescent devices: ZnS [63], ZnS:Mn [64], porous silicon [65], carbon nanotubes [66], etc. Generally speaking, such a composite used as an emitter exhibits a higher conductivity than that of the polymer film alone. Charge carrier mobility is also improved in the composite. However, it is rarely possible to obtain light emission from both materials when using an electroluminescent polymer matrix. On the other hand, the nanocomposite systems show significant porosity [67], which may be a drawback for the fabrication of flexible devices.

Another technique to fabricate micro- and nano-OLEDs was proposed by Granström et al. [68] using poly[3-(4octylphenyl)-2,2′-bithiophene] or PTOPT as an emitter. This technique consists in filling the pores of a microfiltration membrane with PEDOT, which is a conducting polymer. The filled pores then will be used as an electrode with a defined size (10–14 nm). The emitting polymer is spin-coated on top of the membrane-contact structure and a metal is evaporated onto it to form the diode (Fig. 17). Light emission has been obtained from the polymer with dot size of about 100 nm. It should be noted that only 20 to 25% of the diodes fabricated by this technique have worked and the quantum efficiency was estimated to be less than 0.1%. A second diode structure has also been realized by the same authors using the same emitting polymer [69]. The nanosize of the diode in this case was obtained by taking advantage of the phase separation in polymer blends. The size of the emitting polymer islands was experimentally determined by varying the proportion of the components. PTOPT was mixed with PMMA in proportion of 5% PTOPT and 95% PMMA by weight and the solution was spin coated on the PEDOT contact. The size of the nanodiodes is comparable to that obtained by the previous technique.

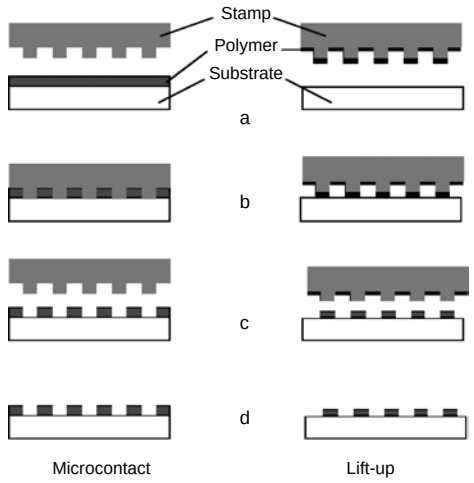


Figure 16. Schematic overview of the microcontact printing techniques: (a) initial setup; (b) conformal contact; (c) stamp removal; (d) patterned polymer. In the microcontact technique, the stamp is dipped in the polymer solution before contact with the substrate. In the lift-up technique, the polymer film is deposited on the substrate before contact with the stamp.

3.4. Ordered Nano-OLEDs

The realization of dispersed nanodiodes has demonstrated that it is possible to fabricate nano-OLEDs, but in practice, a better organization of the devices is needed for displays. Therefore, the conventional techniques used in semiconductor technology are applied to pattern the polymer emitting films. In standard photolithography, the use of solvents to remove the soluble part of the photoresist film may dissolve and swell the organic layer. In addition, UV light can damage the emitter. Methods that do not require lithography have been recently developed for patterning polymer thin films. These methods can be classified into two principal categories: soft lithography and inkjet printing. Other techniques are also available but their applications are limited.

3.4.1. Masking Technique

This technique is the simplest way to produce arrays of OLEDs. A metal mask is put in contact with the substrate and the organic layer is evaporated through the slits of the

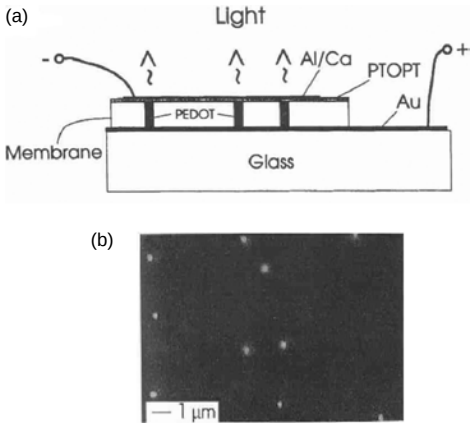


Figure 17. Fabrication of nanodiodes using PTOPT as an emitter, deposited into a polycarbonate membrane: (a) diode structure; (b) photograph of emitting nano-LEDs. Reprinted with permission from [69], M. Granström et al., *Synth. Met.* 76, 141 (1996). © 1996, Elsevier Science.

mask and forms a well defined pattern on the substrate. Generally, there is a close but not direct contact between the mask and the substrate in order to avoid scraping of the organic layer. The formed thin films have high definition geometry but their size is important. Simple pixel arrays of Alq_3 based diodes were fabricated by Jabbour et al. [70] using an electrochemically patterned mask or an electron microscope grid. Mori et al. [71] used a simple sliding mask to fabricate RGB (red green blue) pixels, each of them is 120 μm wide (Fig. 18).

The masking technique was used to fabricate planar OLEDs or surface cells [72]. In this configuration, electrodes were deposited on a polymer substrate with a gap in between using a shadow mask. The cell configuration can be a single gap cell [73] or interdigited one [74, 75]. The gap between the electrodes is a few micrometers wide and the light intensity is proportional to the total length of the electrode. In poly(3-(2′,5′-bis(1″,4″,7″ trioxaoctyl)phenyl)-2,2′-bithiophene) planar based diodes, a turn-on voltage of 6 V was sufficient to run a device with an electrode separation of 10 μm. The major advantage of such diodes over sandwich ones is that they are less sensitive to pinhole formation and may avoid electrical short in operation. However, the width of the LEDs is relatively large, of the order of a few micrometers, and it is not possible at this stage to reduce further the size of devices. Nevertheless, these diodes could be used to realize dot matrix displays.

3.4.2. Standard Lithography

The technique conventionally used in inorganic device manufacture was successfully applied to polymer LEDs by several authors to fabricate microsize diodes. For instance, Lidzey et al. [76] deposited a photoresist layer onto the emitting layer [poly(2,5-dialkoxy-*p*-phenylene vinylene)] and exposed the film to UV radiation through a contact mask. After development and washing, the exposed resist was removed and metal electrode was deposited forming micropatterned arrays of 5 μm diameter devices on a 7 μm pitch. The authors reported a loss of 35% of the photoluminescence efficiency that they attributed to the modification of the polymer during the photoprocessing. To avoid such a problem, a solution was proposed by Nagayama et al. [77] who have used a conventional lithographic method but instead of depositing the resist on the emitting layer, they fabricated arrays of cathode separators directly on the ITO substrate. These arrays of 30 μm width will act as a shadow masks in the vacuum evaporation process of the organic layer and the cathode to form the devices. The principle of the process is shown in Figure 19. This method can be also applied to fabricate polymer based diodes [78]. In this case, the polymer films were cast before patterning the cathode separation barriers. An insulator can be inserted between the emitter and the substrate to prevent edge defects, which may lead to electrical shorts between the cathode and the anode.

Standard lithography is also applied to pattern the anode, especially the ITO substrate. The cathode cannot be patterned the same way because the organic layer would be damaged under the high temperature at which the photoresist is baked. The emitting layer usually dissolves in developers. After deposition of the emitting layer by a conventional technique on the patterned ITO film, the cathode layer is evaporated through a mask to form the diodes. The size of the ITO pattern and that of the metallic electrode define the diode size. This technique was used for fabrication of photodiodes with poly(3-hexylthiophene) and 1-(3-methyoxycarbonyl)propyl-1-phenyl(6,6)C_{61} blend as an active layer [79]. The ITO layer was patterned into strips 450 μm in width and aluminum was deposited on a strip 635 μm wide forming an array of small pixels that could be deposited onto a flexible substrate material. Hosokawa et al. [80] used a similar technique to fabricate RGB multicolor displays, but instead of sizing the diode by fixing the size of the cathode layer, they used the photolithographic process to form ribs on the ITO substrate for the electrode separation. Next, organic layers and cathodes were

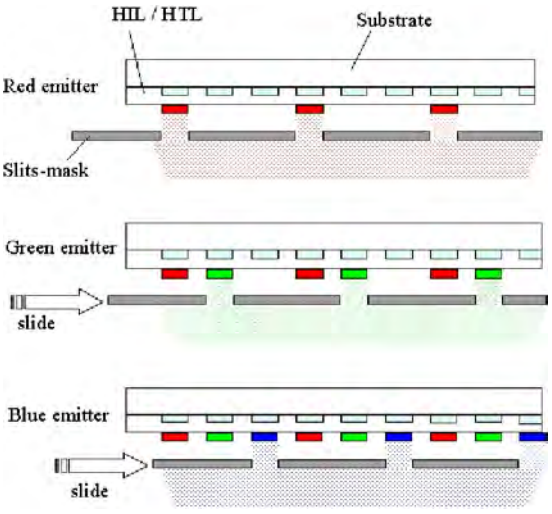


Figure 18. Nano-RGB pixels deposited by a sliding mask technique. Reprinted with permission from [71], K. Mori et al., *Displays* 22, 43 (2001). © 2001, Elsevier Science.

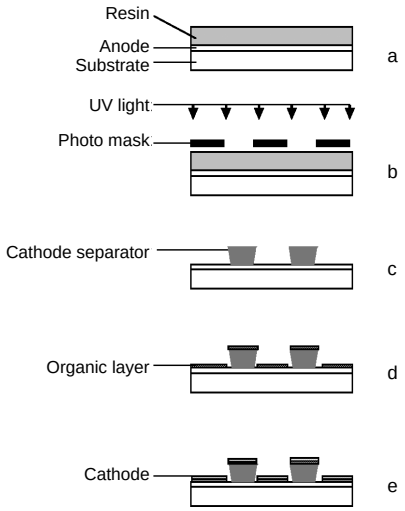


Figure 19. Deposition of nano-OLEDs using the cathode separator technique: (a) initial setup; (b) exposure to UV light; (c) formation of the cathode separators; (d) deposition of the organic layer; (e) deposition of the cathode.

sequentially deposited on the ITO film within the space left between the ribs. Resolution less than 100 μm could be obtained using this technique. Based on the same principle, Tian et al. [81] fabricated Alq_3 devices of 300 μm width using oblique deposition of the cathode to avoid overheating of the emitting layer. The lifetimes of these diodes were found to be significantly improved. An original technique of shadow masking was proposed by Py et al. [82], who fabricated the mask directly on the inorganic substrate with overhanging edges. They used a stack of dense SiO_2, porous SiO_2, and Si_3N_4, deposited on a ITO substrate. To obtain the mask, a photoresist layer was patterned on the stack and was transferred by acid etching. Because of the different rates of dissolution of the insulators, the porous SiO_2 will be undercut and an overhanging profile is formed under the Si_3N_4 layer, which acts as an integrated mask. Rows of 1.54 μm width could be filled with Alq_3 to form an array of OLEDs by this technique.

These techniques are convenient for most of the soluble polymers or evaporated organic materials. For polymers such as PPV, which need a thermal conversion process at high temperature, these techniques are not adapted because of the possible damage of the resist. Visconti et al. [83] used a method based on UV holographic lithography and plasma etching to produce patterning of PPV. This technique consists of making a mask pattern by holographic lithography using an argon laser (363.8 nm) on the resist layer, which is deposited on the polymer film. By rotating the interferometer around the laser axis, the authors could obtain planar arrays of squares, triangles, and diamonds on the resist layer. Next, the patterns are transferred to the PPV film by a reactive ion etching process. By carefully controlling the etching rate, the resist layer was removed and continuous PPV stripes with a width of 200 nm could be formed.

3.4.3. Soft Lithography

Standard lithography has disadvantages in OLED fabrication, mainly because it cannot provide efficient control over the chemistry of patterned organic surfaces. A number of techniques have been developed for fabricating micro- and nanostructures using an elastomeric stamp or mold to transfer the pattern to a substrate. The elastomer stamp is usually fabricated by patterning a silicon substrate with a prepolymer of PDMS. After curing, the stamp is peeled off from the substrate. These techniques, called soft lithography, include μCP, microinjection molding in capillaries (MIMIC), replica molding, and microtransfer molding [84]. To minimize the defects of the pattern, the soft lithography techniques make use of the self-assembly of the molecules. This property of the molecules to be organized into stable configuration is based on the principle that at a final state, their organization is at thermodynamic equilibrium and therefore tends to reject any defect that could be formed. One of the best-known examples of self-assembly is the self-assembled monolayer, which is formed by self-organization of functionalized organic molecules on an appropriate surface.

Soft lithography techniques have been used by Grandlund et al. [85] to fabricate arrays of OLEDs. The arrays were formed by patterning the conducting polymer PEDOT/PSS layer deposited on an ITO substrate using the μCP technique. After deposition of the active polymer layer [poly(3-(2-butyloxy-5-octylphenyl))], they used the MIMIC technique to pattern the polymer film, and finally, a layer of aluminium was deposited over the active layer to form arrays of OLEDs. The size of each diode was 100×100 μm^2, which has a similar performance to that of standard devices.

A similar technique has been used by Nüesch et al. [86] to pattern the ITO anode (Fig. 20). A elastomeric PDMS stamp was fabricated to the desired pattern and then immersed in a 3×10^{-3} M solution of (tetrabutylammonium) hydroxide in ethanol or water. The stamp is then pressed onto the ITO substrate. The base will react with the conducting oxide film, resulting in the formation of a solid ionic double layer at the oxide surface and leading to a change in its work function [87]. The devices were obtained by depositing a TPD/Alq_3 double layer onto the ITO film and the cathode LiF/Al was finally evaporated to form the diodes. The patterned ionic layer will inhibit the positive charge injection from the anode and only the organic film corresponding to the unpatterned parts of the ITO substrate will emit light. A luminance of about 100 cd/m^2 was obtained with a driving bias of 6.5 V for such a diode. This method has the advantage in that it will avoid handling of the organic film, which may cause damage to the devices.

Arrays of OLEDs with a period of 200 nm were reported by Wang et al. [88], who used nanoimprint lithography at 150 °C in vacuum. The emitting layer was Alq_3 doped with 2 wt% DCMII dye molecules, which has high luminescence efficiency. The patterning process in vacuum was found to preserve the optical property of the organic layer by preventing the residual oxygen from contaminating it. Arrays of about 400 nm width were realized by Nyberg et al. [89] using poly(3-(2′-methoxy-5′-octylphenyl)thiophene) as an emitting layer. Making use of this method, an insulating layer was use to separate the cathode from the anode, the pattern being obtained on the insulating layer and the cathode. The emitting film was deposited in the space left between two

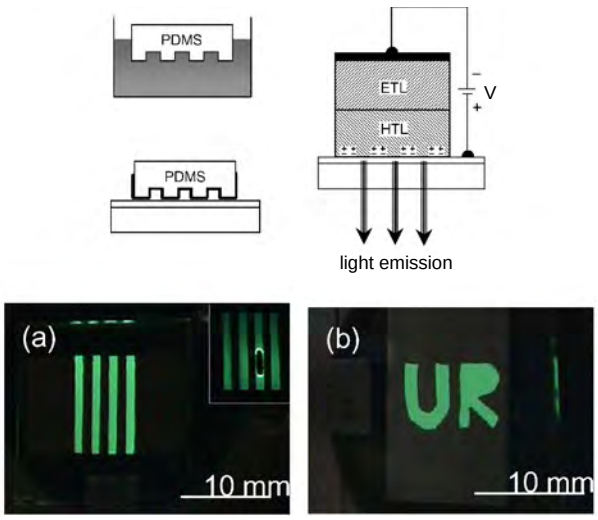


Figure 20. Fabrication of nano-OLEDs with Alq_3 as an emitter, using a soft-lithography technique to pattern the ITO substrates. Reprinted with permission from [86], F. Nüesch et al., *Appl. Phys. Lett.* 75, 1799 (1999). © 1999, American Institute of Physics.

successive arrays and the diodes have a quantum efficiency similar to standard sandwich devices (0.002%).

It is also possible to pattern the metal cathode layer directly using soft lithography. This technique makes use of silicon stamp to pattern the cathode layer by applying pressure onto the surface of the diodes (cold-welding process [90]). The stamp was previously patterned with raised areas in the shape of the regions on the electrode that were to be removed. When this was done, the left areas on the substrate can be used as arrays of LEDs. The active material used in this fabrication was Alq_3 with an electron transport layer made of 4,4′-bis(*N*-naphthyl)-(*N*-phenyl-amino)biphenyl. The measured performance on 1 mm diameter diodes is comparable to those obtained in Alq_3 based LEDs and no apparent degradation of the emitting layer was observed during the cathode patterning process.

The different soft lithography techniques have demonstrated that they can be applied to the OLED technology without producing significant damage to the organic layer and the performances of the nano- and microdiodes are generally comparable to those of standard ones. Devices of small size and high definition could be obtained by these techniques as compared to standard lithography. Recent investigations have demonstrated that patterns with 10 nm features could be prepared [91]. However, it should be noticed that industrial production requires rapid and reproducible patterning, and parallel techniques of fabricating devices are more adequate for manufacturing nanostructures. Soft lithography methods are well adapted to serial fabrication of devices in laboratory research work or lithographic masks.

3.4.4. Inkjet Printing

This technique is particularly adapted for solution processable materials and appears to be the most interesting for fabrication of devices with highly resolved patterns [92, 93]. The emitting layer can be produced by two different methods. For the first technique, the film is directly deposited by printing on the substrate. For the second, the polymer is preliminarily spin-coated onto the substrate. Then a solution containing the dopant is dropped over the films, allowing the dye diffusion into the host layer.

Inkjet printing can be performed in a continuous way or in a drop-on-demand (or impulse) mode. The difference between the two modes is in the pattern size and the printing accuracy, the drop-on-demand mode allowing smaller size patterning and a better accuracy. This process can be performed by using either a bubble jet or a piezoelectric technique. In the bubble jet technique, the ink droplet is pulled out of the nozzle by the pressure of a vapor bubble that is locally formed by thermal heating. This technique therefore needs water as a solvent for the material. In the piezoelectric technique, the ink droplet is mechanically pumped out of the nozzle by applying a voltage to a piezoelectric crystal. This technique can be applied to polymers dissolved in organic solvents. Notice that many nozzles can operate in parallel to achieve high speed printing, resulting in a low cost production of devices.

Most of the materials used for inkjet printing belong to the polycarbazole (PVK) and the phenylene vinylene families. Light emitting diodes with low turn-on voltage were fabricated by Hebner et al. [94] using PVK doped with coumarin dyes. Soluble derivatives of PPV such as MEH-PPV [95] and poly[5-methoxy-(2-propanoxysulfonide)-1,4-phenylene vinylene] [96] also used successfully make use of the printing technique. Figure 21 shows the polymer light emitting fabrication process and a logo created by inkjet printing using MEH-PPV as an active layer.

The size and the quality of the printed pattern depend on several factors. First, the size of standard nozzles is in the range of 20–30 μm, which produces at best dots of about 25 μm in size [97]. Second, the resolution of the pattern depends greatly on the viscosity of the ink and its ability to bind to the substrate. Generally, a low viscosity is needed to prevent rapid drying of ink and clogging of nozzles. Therefore, additives should be incorporated to the ink when using soluble dyes for color emission. On the other hand, ink droplets should not spread on the substrate before they are absorbed and dried (few seconds to several minutes). This condition implies that the ink viscosity should be sufficiently high and the surface of the substrate should be rough enough to limit spreading of the ink and blurring of the image. Finally, the printed pattern should be of a uniform thickness to avoid possible electrical shorts and the uniformity of the layer depends greatly on the quality of the material used. Indeed, the pinhole formation in printed patterns is usually observed and to improve the quality of the emitting devices, additional layers are used. These layers of course have their usual function but are useful to prevent shorts. For instance, PEDOT doped with PSS was used

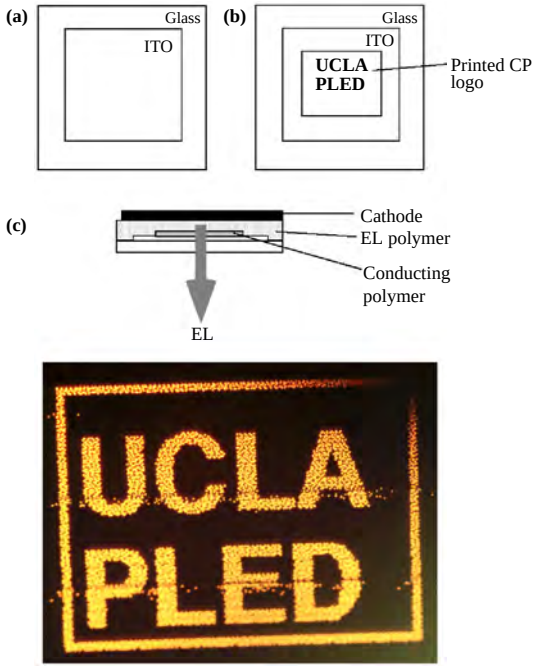


Figure 21. Fabrication of nano-OLEDs with MEH-PPV as an emitter, using an inkjet printing technique to design a logo. Reprinted with permission from [95], J. Bharathan and Y. Yang, *Appl. Phys. Lett.* 72, 2660 (1998). © 1998, American Institute of Physics.

as a hole transport layer in printed OLEDs with polyfluorene and PPV as emitters [98]. Improved techniques and materials [regioregular poly(3-hexylthiophene-2,5 dyil)] have demonstrated that good quality patterns could be realized without buffer layers [99].

The inkjet technique has several undeniable advantages in that a pattern of any size and/or shape can be reproduced directly on the substrate. Problems encountered in the technique include uniform thickness of the printed layer, pinhole formation, and viscosity of the solution. Even with these problems, the technique appears to be one of the most promising for patterning organic devices in the future.

3.4.5. Special Techniques

Besides the techniques currently used for fabrication of small devices previously mentioned, several others have been also carried out with success to pattern organic diodes.

First, a variation of the lithography technique was used by Noach et al. [100], who used a 193 nm excimer laser to ablate first the ITO substrate through a mask. After deposition of the polymer film and the metal electrode, the grid was then rotated orthogonally to the direction of the ITO strips and the laser was used to ablate the metal layer and partially the polymer. Arrays of 20 μm square pixels were reported with improved luminance as compared to standard thin film LEDs. This technique requires, however, careful control of the laser operation to obtain the desired results.

For printing techniques, an alternative method has been applied by Pardo et al. [101] to fabricate logos using MEH-PPV as an emitter (Fig. 22). This technique, called screen-printing, consists of using a nylon fabric as a screen through which the ink (solution) is deposited on an ITO substrate. The screen is in direct contact with the ITO film or close to it ($\approx$1 mm). The shape of the screen can be defined by using an epoxy to hide parts of the fabric that will not be printed. Multilayer diodes can then be formed on the screen by conventional deposition techniques. The technique is relatively simple and allows a high printing speed with a good resolution, below 100 μm. Birnstock et al. [102] reported that very a low turn-on voltage (3 V) was obtained in screen-printed OLEDs using a polyfluorene derivative, without pattern caused by the fabric.

More specific techniques have been recently developed to obtain nanosize OLEDs. These techniques can be divided into two main categories: the first one uses optical means to obtain a light emission from the organic material and the second is a special preparation of the substrate to pattern the devices.

The nanodevices can be patterned by light [103]. The technique used consists of fabricating a standard diode and then irradiating it by an intense light beam (laser) through the transparent substrate. An alternative voltage is simultaneously applied to the device during the irradiation (several minutes). It was observed in Alq_3 based diodes that electroluminescence was induced by the light beam irradiation and a memory effect was obtained: the device continued to emit light during a week. This observation suggests that the illumination facilitates the charge injection and their transport through the device, leading to the formation of excitons. However, the memory effect induced by the light is not well understood. This technique allows patterning the diodes by moving the light beam on the surface of the substrate.

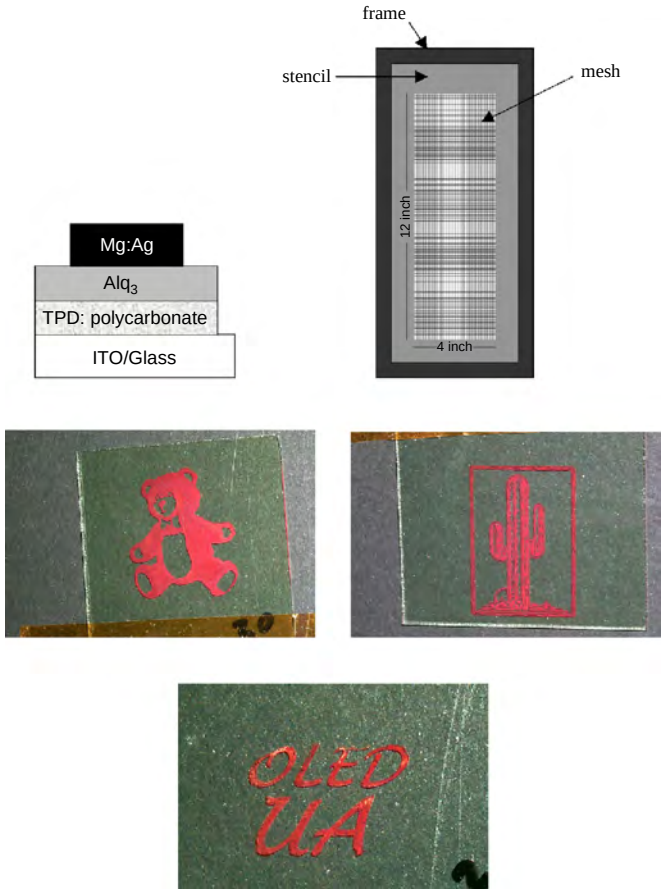


Figure 22. Fabrication of nano-OLEDs with MEH-PPV as an emitter, using a screen printing technique to design a logo. Reprinted with permission from [101], D. A. Pardo et al., *Adv. Mater.* 12, 1249 (2000). © 2000, Wiley-VCH.

A variation of the previous technique was used by Micheletto et al. [104] with a laser light of a modified scanning near field optical microscope setup. The device studied was a ITO/poly-3(2-(5-chlorobenzotriazolo)-ethyl)thiophene/Al diode. The cathode is the tip of the microscope, which acts as a nanosize electrode. It allows stimulation of very small areas of the active polymer and locally complete LEDs are formed. By this technique, it was possible to study diodes of size about 100 nm and significant differences in the light emission of these diodes with that of standard devices were observed. It should be noticed that the technique does not allow one to obtain permanent devices but may be helpful for studying properties of small size diodes. Based on the same principle, other optical instruments were modified and used to study very small areas of OLEDs. For instance, diaminotofluorbiphenyl based diodes were studied with a scanning tunneling microscope [105] and 1,3,4-oxadiazole moiety modified alkoxy ring substituted poly(*p*-phenylene vinylene) LEDs by atomic force microscopy [106].

Highly ordered OLED heterostructures were fabricated by Marks et al. [107] using a special preparation of the

substrate. This technique is called nanosphere lithography (NSL). It consists of making a special mask deposited directly on the substrate (Fig. 23). First, a self-assembled layer of triarylamine was deposited on the ITO substrate in order to obtain a smooth and uniform surface. Then a layer of hexagonally close-packed carboxy-terminated polystyrene beads was formed on the obtained surface, playing the role of a mask. Diodes were fabricated by evaporation of Alq_3 as an emitting layer and TPD as a hole transporting layer. Regular pattern prismatic posts with 90 nm sides were obtained, which emit a light similar to that of macroscopic diodes without a significant shift of the EL maximum. However, an overall light intensity was found to be two orders lower than that of standard diodes, working in similar bias conditions. This technique provides very regular small size diodes.

3.4.6. Concluding Remarks

Many techniques have been invented or borrowed from the inorganic material technology to fabricate nano-OLEDs. The size of the devices ranges from several tens of nanometers to several micrometers. Generally, the performance of small size diodes is similar to that of conventional ones. However, it should be noted that the size of the devices can influence the physical process involved in transport and light emission of the diodes. These particular aspects will be examined. Moreover, although most of techniques used give reliable and reproducible results, only a few of them are suitable for industrial fabrication, which requires low cost, mass production, and high performance devices. Finally the technique used should also suit most of the organic materials to allow the fabrication of multiplayer and multicolor devices, which are necessary for the realization of displays.

3.5. Applications: Matrix Displays

Applications of OLEDs in the field of displays can be divided into two types: active matrix and passive matrix. These displays differ from each other in their structure and in the controlling process. A passive matrix has a simpler structure than an active matrix, but its control is less efficient than that of an active one.

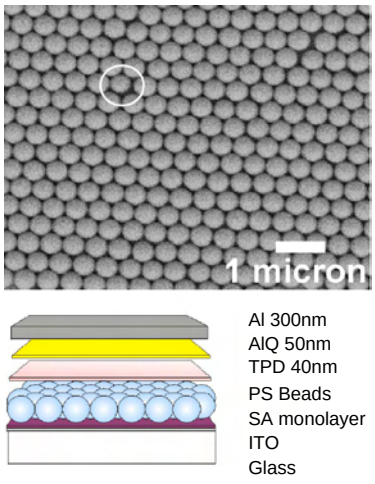


Figure 23. Schematic overview of the NSL. Reprinted with permission from [107], T. Marks et al., *Synth. Met.* 127, 29 (2002). © 2002, Elsevier Science.

3.5.1. Passive Matrix Displays

Principle Passive matrix displays consist of an array of picture elements or pixels deposited on a substrate. The pixels are patterned by one of the techniques described in previously. They are formed at intersections of rows (X horizontal) and columns (Y vertical) lines. To switch on a particular pixel (X, Y), electrical signals are applied to the X and Y lines, with a positive bias to the anode and a negative bias to the cathode. Each pixel is composed of a light emitting diode, whose color is designed by the choice of the emitting material. For a white pixel, three red, green, and blue subpixels are necessary. The brightness of a pixel is controlled by setting the current through each individual diode to the desired value.

Current Cross-Talk The circuit scheme of a OLED passive matrix is shown in Figure 24. Each pixel is switched on when a bias higher than its turn-on voltage is applied. However, because of the electrical connections between the diodes, which are connected to the same row or to the same column, parallel conduction paths occur when a selected pixel is switched on. In other words, a dc current cross-talk exists and may cause unwanted light in the matrix. There are several sources of cross-talk [108]: (i) the display resolution which creates more conduction paths when it is high, (ii) the diode rectification ratio, which reduces the cross-talk when it is high, and (iii) the diode reverse leakage current, which increases the stray light when it is high. Circuit simulation performed by Braun et al. [109] indicates that a current driver is more suitable than a voltage driver to address pixels. Furthermore, the electrode resistance, the diode leakage current, and the location of faulty pixels will strongly influence the image uniformity and the power consumption of the display. The cross-talk problem can be minimized by using a carbon underlayer placed between the ITO and the hole transport layer [110]. This layer insulates the electrodes and thus lowers the leakage current, preventing the occurrence of cross-talk.

Design Aspects for Displays In a passive matrix display, each line of the picture is addressed sequentially for a short time. The forward biased pixels will emit light while those nonaddressed are kept at zero or reverse bias. To avoid possible flicker effects, the scanning period should

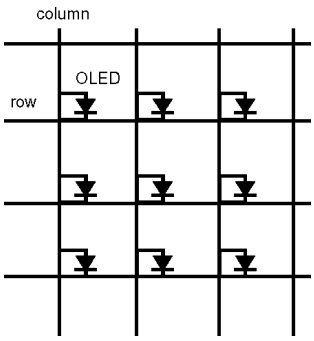


Figure 24. Schematic circuit of an OLED passive matrix.

be smaller than $1/XV_r$ sec, where V_r is the video rate (V_r = 25 images sec^{-1} for PAL standard). However, because the scanning period is very short, the emitted light intensity should be high enough to allow the eye to see a picture with a normal brightness. This condition implies that the average luminance perceived on the screen should be multiplied by a factor equal to the number of lines (X) to obtain the real luminance of the diodes. Beside this consideration, other factors should be taken into account in the calculation of the device luminance. For a monochrome display, one should compensate the loss in emission due to the use of a contrast filter and to the faulty pixels [111]. For instance, to obtain an average of 100 cd/cm^2 in a display of 120 lines, with 20% of faulty pixels, the required luminance of the diodes is 37,500 cd/cm^2. For a color display, a factor of 12 higher peak luminance would be needed because each pixel is composed of three RGB subpixels that have a reduced area, a reduced scanning time, and an increased number of lines. Of course, such a high luminance of the devices will require a high power consumption leading to a high power dissipation by Joule effect that should be reduced, for instance, by dividing the anode line into different sectors, each driven by independent synchronous circuitries.

Displays Passive matrix displays were fabricated and studied by using polymer or organic materials as an emitter. They utilize a simple structure, which is suited for low cost, low information applications such as alphanumeric displays. The techniques for the realization of the devices are different depending on the material used. For polymers, standard and soft lithography are well adapted for making small size pixels. Inkjet printing is also convenient for polymers, especially soluble ones, which can be deposited into determined size pixels. For organic materials, vacuum evaporation with the use of a special masking technique is often adopted.

Monochrome displays are currently fabricated by many companies. The first realization was the seven-segment prototype fabricated by Cambridge Display Technology using a PPV derivative. In 1997, Pioneer launched their passive matrix displays of 256 × 64 dots for automotive applications, based on Alq_3 diodes. In 2000, Motorola fabricated a cellular phone screen with passive matrix OLED displays (Timeport P8767) from Pioneer. Researchers from Huyndai LCD presented a 2.4 in. blue monochrome display for mobile application, with 160 × 160 pixel resolution and an average brightness of 30 cd m^{-2} [112]. High resolution and full color displays have been developed in recent research. In 1997, Hosokawa et al. [80] fabricated an organic display of 5 in. with 320 × 240 pixels (256 gray scale) which was followed by a full color display of 10 in. with VGA resolution [113]. The pixel surface was 330 × 320 μm^2. Pure RGB color passive matrix displays were developed by Pioneer with 320 × 240 pixels of size of 330 × 330 μm^2 and a luminance of 150 cd m^{-2} [114, 115] using a high accuracy moving system mask to deposit RGB emitters.

Despite the progress made in this field, the passive matrix displays need to be driven under pulsed conditions at high current density, which is not suitable for high resolution and high information content applications such as videos or graphics. For such applications, active matrix displays are used.

3.5.2. Active Matrix Displays

Principle In an active matrix display, each pixel element can be addressed independently via the associated thin film transistors (TFTs) and capacitors in the electronic backplane (Fig. 25). There are two TFTs employed for driving an OLED: a drive transistor and an address transistor. To address a particular pixel, the proper row is switched on and then a charge is sent down to the correct column. A storage capacitor at the designated pixel receives this charge and is able to hold the charge until the next refresher cycle. The addressed transistor acts as a switch for the drive transistor by holding a voltage on the gate of the latter. The OLED pixel has to be biased with a constant current and is thus connected to the drain of the drive TFT mounted in series with the diode. The current can be supplied to the diode during the whole frame time and the current density is much smaller than that needed to drive a passive matrix display. Other advantages of the active matrix are no cross-talk problems, continuous excitation, low power consumption, and integrated drive electronics.

Design Aspects for Displays Active matrix displays are currently fabricated using amorphous or polysilicon deposited on glass or plastic substrates although conjugated polymer has been also proposed to drive OLEDs [116, 117]. Poly-Si can be deposited at low temperature but has several drawbacks, including low electron mobility and variation in electrical properties, leading to difficult control of the current density at each pixel. Several techniques can be used to improve the performance of Si TFTs, including low temperature crystallization or excimer laser annealing [118, 119]. However, these techniques require drastic control of the substrate, which make them difficult to put into practice. A four-transistor circuit has been proposed to remedy these drawbacks [120, 121] (Fig. 26). This circuit makes use of the gate voltage of a TFT to control the data current level at a given value, regardless the TFT voltage threshold, resulting in constant output current density which switches on the pixel. A luminance uniformity and an improved performance at low luminance levels were obtained in the 16 × 16 pixel display prepared for this study. It should be noted that *p*-type TFTs are more adapted for driving OLEDs due to the fabrication process used for the organic diodes. In fact, the use of an *n*-type TFT would require the anode (ITO) of the diode to be deposited on the organic layer, which can damage the emitter.

Displays Although the feasibility of OLED active matrix displays was demonstrated a few years ago [122, 123], it is known that one TFT pixel electrode configuration cannot be used because of the continuous excitation during the frame period. Therefore, a production of TFT substrates is needed

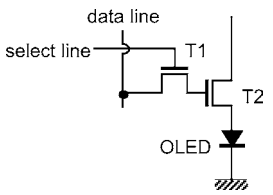


Figure 25. Schematic circuit of an OLED active matrix.

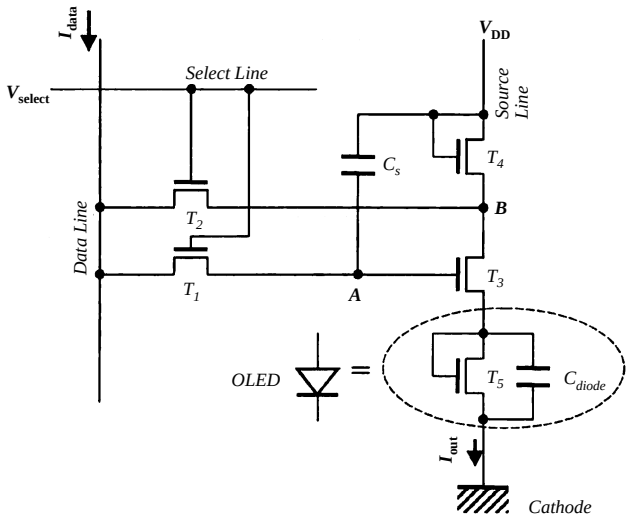


Figure 26. Schematic circuit of a four-TFT OLED active matrix. Reprinted with permission from [121], Y. He et al., *Jpn. J. Appl. Phys.* 40, 1199 (2001). © 2001, Institute of Pure and Applied Physics.

to develop active matrix displays. Using organic materials, Pioneer Corp. has fabricated a 2 in. device with 176 × 192 dots targeted at mobile telephones and a 3 in. device with 320 × 320 dots targeted at personal digital assistants (PDAs) [124]. A 160 × 160 pixel display is usually considered sufficient for a square PDA screen. Dawson et al. [125] described a prototype of QVGA screen (320 × 240 pixels) using a four-transistor autozeroing pixel driver. An alternative approach has been proposed by eMagin by fabricating OLEDs on silicon microdisplays [126]. Instead of low temperature polysilicon or amorphous silicon, the driving circuit is directly integrated within the silicon wafer using CMOS technology. Full color SVGA (852 × 3 × 600 pixels) displays have been realized with a high brightness (>100 cd m^2) and a high lifetime (>10,000 hrs) and can be used in applications such as wearable PCs, cell phones, and digital cameras. Another technique for driving the pixels has been proposed by Seo et al. [127] using a simple circuit (capacitance and rectifying diode) and providing a good control of the gray scale. Attempts to fabricate active matrix displays on flexible substrates have been performed by several research teams but no prototype has yet been proposed. These realizations need a low temperature process to implant TFT on plastic substrates (150 °C for polyethylene terephthalate and 200 °C for polyethersulphone).

3.6. Physics of Nanosize Devices

It is tempting to believe that the background we know from conventional solid-state physics can be applied without difficulty to nanodevices. However, the size of the objects involved in nanotechnology is such that the structure may significantly influence the physical processes that occur in devices [128]. In the field of OLEDs, several aspects can be considered.

First, from the electrical point of view, because of the small size of the devices, the current density would be very large. There will be a problem of energy dissipation (intensity and mechanisms) inside the structure which at this time is not clearly understood. On the other hand, the charge injection processes are believed to dominate in molecular OLEDs contrarily to molecular junctions [129], in which injection cannot occur because of the minute thickness involved (a few nm). Nevertheless, it is known that the injected charges modify the electrostatic potentials at the interface and chemical interactions can occur between the organic material and the electrode. The effects of these processes on the transport can be enhanced in very small size devices and this still needs to be elucidated [130, 131].

Second, from the optical point of view, the size of the devices may influence the emission properties of the organic materials. In the nanometer range, quantum confinement effects are expected to blueshift the optical absorption edge with decreasing material size. Small Si crystallites of 5 nm are found to emit intense blue light [132]. The optical response and charge transport in electronic polymers can be drastically changed by a small variation in local order in the materials, for instance by modifying the substituent group [133]. Experiments performed in Alq_3 of 10–15 nm sizes clearly show that the PL spectra of the material are blueshifted and the shift direction depends strongly on the size of the sample [134]. In large size samples, the formation of dimers or multimers is probable and their strong interaction results in an increase of the gap between the excited states, producing a redshift of the PL. In smaller size samples, the organic molecules would resemble monomers and the PL spectra are blueshifted.

The effects of the intermolecular interactions on the photophysical and transport properties of organic semiconductors can also be investigated through the nanoengineering of the materials, by modifying the structure of conjugated dendrimers [135] so as to understand the physical mechanisms involved within the nanostructures.

The miniaturization of the devices requires much research on new physical issues, which include, among others, the hot electron effect, the phonon effect, the quantum noise effect, the purity effect, and those effects already mentioned.

4. CONCLUSION

With the intensive research of recent years, the development of organic devices has attained a level which allows industrial applications to be commercialized, especially in the display field. Electronic companies have started to fabricate small screens for cell phones and handheld computers which compete with those using conventional techniques. Organic materials have several undeniable advantages over classical semiconductors. In particular, their very high luminance is a definite asset in display applications. These materials are also well adapted to nanotechnology, resulting from the ease of the deposition technique, which allows any shape and size of the devices to be manufactured. Not only can conventional methods developed for semiconductor technology be applied to polymers and organic materials, but also new ones (inkjet, screen printing) have been invented to fabricate nanosize devices. These techniques benefit from the unique physical properties of the materials (solution, flexibility) that are not evident in conventional semiconductors.

High-resolution displays need very small size pixels. Technically speaking, organic diodes of submicrometer size can already be fabricated with an acceptable performance. However, the question of tolerance still remains a problem. Fluctuations in size and in structural properties of the emitting layer may severely affect the quality and performance of the diodes and need to be limited to an acceptable level. Furthermore, under normal working conditions, the tiny devices have to dissipate a considerable amount of energy in regard to their size. The stability of the emitting materials is therefore the key for the fabrication of reliable displays. For the commercialization of these products, mass production would need a simple, low cost, and reproducible process for fabricating the emitting layer. At present, inkjet or screen-printing using polymers appears to be the best technique for making emitting devices.

Another concern in the field of nano-OLEDs is the interconnections. The devices in matrix displays need to be driven by appropriate circuits that require reliable electrical contacts. Although this problem can be solved by techniques used in conventional electronics, it has been proved that a number of faulty pixels exist because of failure in electrical connection.

The development of organic nanodiodes has been accompanied by new physical issues that have to be further investigated. In addition to the power dissipation in these devices, the electrical transport process, the light emission, and the interface between the emitter and the electrodes have a new and different aspect as compared to conventional and standard diodes. Not only does the size influence these properties, but also new phenomena occur, especially in the optical processes, which are not yet clearly understood.

According to many observers, the recent progress in microelectronics technology has confirmed Moore's law and this trend will continue to obey the law until the year 2010. Projections will then have to be reformulated. Further research and development appear necessary to replace the conventional semiconductors in the near future. Organic materials are good candidates, at least in the field of optoelectronics, due to their excellent emission properties. As regards nanotechnology, they are quite suitable to the processes and technologies used for the fabrication of nanodevices. There is great expectancy that these materials will have a bright and brilliant future in a new nanoelectronics era, when the ultimate limitations of the component size will be reached.

GLOSSARY

Electroluminescence Emission of light from a material when excited by an electrical mean (applied voltage). The light is generated by recombinations of excitons.

Exciton A pair of electron-hole generated in a material by an external source (light, electrical mean) and bound together by electrostatic attraction. An exciton can be a singlet (the electron and hole have spins in opposite directions) or a triplet (the electron and hole have spins in the same direction).

Lithography Printing technique that puts detailed patterns on a flat surface of a substrate. In the classical photolithographic process, an image of the desired pattern is projected onto the surface of the substrate, which is coated by a thin layer photosensitive resist. The bright parts of the resist film become soluble by a chemical reaction with the light, whereas the dark parts remain insoluble. After contact with a developer solution, the soluble part of the resist is removed. The pattern is transferred onto the substrate by a chemical etching of its surface, which preserves the protected parts by the resist film. Other lithographic techniques include electron-beam lithography, soft X-ray lithography, atomic-beam holography.

Nanoimprint Printing technique using a stamp of nanometer size to transfer a polymer film onto a solid substrate surface. The formed polymer film is then used either as an active layer in devices or as a mask in a conventional lithographic process.

Pixel A picture element. In OLED technology, a pixel is determined by the part of the emitting layer where the cathode and anode overlap. A high display resolution needs small size and closely spaced pixels.

REFERENCES

1. M. Pope, H. Kallmann, and W. G. Schneider, *J. Chem. Phys.* 3, 2042 (1963).
2. C. W. Tang and S. A. Van Slyke, *Appl. Phys. Lett.* 51, 913 (1987).
3. J. H. Burroughes, D. D. C. Bradley, A. R. Brown, R. N. Marks, K. Mackay, R. H. Friend, P. L. Burns, and A. B. Holmes, *Nature* 374, 539 (1990).
4. M. Pope, H. P. Kallmann, and P. Magnante, *J. Chem. Phys.* 38, 2042 (1963).
5. W. Helfrich and W. G. Schneider, *Phys. Rev. Lett.* 14, 229 (1965).
6. H. Ishii, K. Sugiyama, E. Ito, and K. Seki, *Adv. Mater.* 11, 605 (1999).
7. C. Adachi, S. Tokito, T. Tsutsui, and S. Saito, *Jpn. J. Appl. Phys.* 27, L269 (1988).
8. H. Meyer, D. Haarer, H. Naarman, and H. H. Hothold, *Phys. Rev. B* 52, 2587 (1995).
9. K. Pichler, C. P. Jarrett, R. H. Friend, B. Ratier, and A. Moliton, *J. Appl. Phys.* 77, 3523 (1995).
10. P. S. Davids, S. M. Kogan, I. D. Parker, and D. L. Smith, *Appl. Phys. Lett.* 69, 2270 (1996).
11. P. Mark and W. Helfrich, *J. Appl. Phys.* 33, 205 (1962).
12. D. M. Pai, *J. Chem. Phys.* 52, 2285 (1970).
13. W. D. Gill, *J. Appl. Phys.* 43, 5033 (1972).
14. H. Bassler, *Phys. Status Solidi B* 175, 15 (1993).
15. Yu. N. Gartstein and E. M. Conwell, *Chem. Phys. Lett.* 245, 351 (1995).
16. S. V. Novikov, D. H. Dunlap, V. M. Kenkre, P. E. Parris, and A. V. Vannikov, *Phys. Rev. Lett.* 81, 4472 (1998).
17. U. Albrecht and H. Bässler, *Chem. Phys.* 199, 207 (1995).
18. P. Langevin, *Ann. Chim. Phys.* 28, 433 (1903).
19. "Primary Phoroexcitaions in Conjugated Polymers: Molecular Exciton versus Semiconductor Band Model" (N. S. Sariciftci, Ed.). World Scientific, Singapore, 1997.
20. A. S. Davydov, "Theory of Molecular Exciton." Plenum Press, New York, 1971.
21. E. M. Conwell, *Synth. Met.* 83, 101 (1996).
22. Y. Cao, I. D. Parker, G. Yu, C. Zhang, and A. J. Heeger, *Nature* 397, 414 (1999).
23. Z. Shuai, D. Beljonne, R. J. Silbey, and J. L. Brédas, *Phys. Rev. Lett.* 84, 131 (2000).

24. M. A. Baldo, D. F. OBrien, Y. You, A. Shoustikov, S. Sibley, M. E. Thompson, and S. R. Forrest, *Nature* 395, 151 (1998).
25. M. A. Baldo, M. E. Thompson, and S. R. Forrest, *Nature* 403, 750 (2000).
26. J. H. Burroughes, D. D. C. Bradley, A. R. Brown, R. N. Marks, K. Mackay, R. H. Friend, P. L. Burns, and A. B. Holmes, *Nature* 341, 539 (1990).
27. C. W. Tang, S. A. van Slyke, and C. H. Chen, *J. Appl. Phys.* 65, 3610 (1989).
28. H. Shirakawa, E. J. Louis, A. G. MacDiarmid, C. K. Chiang, and A. Heeger, *J. Chem. Soc. Chem. Commun.* 578 (1977).
29. J. Cornil, D. Beljonne, D. A. Dossantos, Z. Shuai, and J. L. Brédas, *Synth. Met.* 78, 209 (1996).
30. Y. Shirota, *J. Mater. Chem.* 10, 1 (2000).
31. J. Q. Zhao, S. J. Xie, S. H. Han, Z. W. Yang, L. N. Ye, and T. L. Yang, *Synth. Met.* 114, 251 (2000).
32. J. S. Kim, M. Granström, R. H. Friend, N. Johansson, W. R. Salaneck, R. Daik, W. J. Feast, and F. Cacialli, *J. Appl. Phys.* 84, 6859 (1998).
33. G. Gustafsson, Y. Cao, G. M. Treacy, F. Klavetter, N. Colaneri, and A. J. Heeger, *Nature* 357, 477 (1992).
34. T. M. Brown, J. S. Kim, R. H. Friend, F. Cacialli, R. Daik, and W. J. Feast, *Appl. Phys. Lett.* 75, 1679 (1999).
35. T. J. Gardner, C. D. Frisbie, and M. S. Wrighton, *J. Am. Chem. Soc.* 117, 6927 (1995).
36. F. Nüesch, F. Rotzinger, L. Siahmed, and L. Zuppiroli, *Chem. Phys. Lett.* 288, 861 (1998).
37. S. F. J. Appleyard, S. R. Day, R. D. Pickford, and M. R. Willis, *J. Mater. Chem.* 10, 169 (2000).
38. L. S. Hung, C. W. Tang, and M. G. Mason, *Appl. Phys. Lett.* 70, 152 (1997).
39. R. Schlaf, B. A. Parkinson, P. A. Lee, K. W. Nebesny, G. Jabbour, B. Kippelen, N. Peyghambarian, and N. R. Armstrong, *J. Appl. Phys.* 84, 6729 (1998).
40. M. Y. Hwang, M. Y. Hua, and S. A. Chen, *Polymer* 40, 3233 (1999).
41. G. Grem, G. Leditzky, B. Ullrich, and G. Leising, *Adv. Mater.* 4, 36 (1992).
42. G. Yu, Y. Q. Liu, X. Wu, D. B. Zhu, H. Y. Li, L. P. Jin, and M. Z. Wang, *Chem. Mater.* 12, 2537 (2000).
43. M. Koppe, H. Neugebauer, and N. S. Sariciftci, *Mol. Cryst. Liq. Cryst.* 385, 101 (2002).
44. R. J. Curry and W. P. Gillin, *Appl. Phys. Lett.* 75, 1380 (1999).
45. R. G. Sun, Y. Z. Wang, Q. B. Zheng, H. J. Zhang, and A. J. Epstein, *J. Appl. Phys.* 87, 7589 (2000).
46. B. J. Jung, C. B. Yoon, H. K. Shim, L. M. Do, and T. Zyung, *Adv. Funct. Mater.* 11, 430 (2001).
47. J. Kido, K. Hongawa, K. Okuyama, and K. Nagai, *Appl. Phys. Lett.* 64, 815 (1994).
48. B. W. d'Andrade, M. E. Thompson, and S. R. Forrest, *Adv. Mater.* 14, 147 (2001).
49. N. C. Greenham, R. H. Friend, and D. D. C. Bradley, *Adv. Mater.* 6, 491 (1994).
50. H. Aziz, Z. D. Popovic, N.-X. Hu, A.-M. Hor, and G. Xu, *Science* 283, 1900 (1999).
51. D. Kolosov, D. S. English, V. Bulovic, P. F. Barbara, S. R. Forrest, and M. E. Thompson, *J. Appl. Phys.* 90, 3242 (2001).
52. P. F. Smith, P. Gerroir, S. Xie, A. M. Hor, Z. Popovic, and M. L. Hair, *Langmuir* 14, 5946 (1998).
53. Y. F. Liew, H. Aziz, N. X. Hu, G. Xu, and Z. Popovic, *Appl. Phys. Lett.* 77, 2650 (2000).
54. L. R. Harriott, *Proc. IEEE* 89, 366 (2001).
55. M. Austin and S. Y. Chou, *J. Vac. Sci. Technol. B* 20, 665 (2002).
56. S. Y. Chou, P. R. Kraus, and P. J. Renstrom, *Science* 272, 85 (1996).
57. H. Li and W. T. S. Huck, *Curr. Opin. Solid State Mater. Sci.* 6, 3 (2002).
58. B. Faircloth, H. Rhors, R. Tiberio, R. Ruolff, and R. R. Krchnavek, *J. Vac. Sci. Technol. B* 18, 1866 (2000).
59. E. Delamarche, H. Schmid, A. Bietsch, N. B. Larsen, H. Rothuizen, B. Michel, and H. A. Buiebuyck, *J. Phys. Chem. B* 102, 3324 (1998).
60. E. Schäffer, T. Thurn-Albrecht, T. P. Rusel, and U. Steiner, *Nature* 403, 874 (2001).
61. S. Guha, R. A. Haight, N. A. Bojarczuk, and D. W. Kisker, *Appl. Phys. Lett.* 82, 4126 (1997).
62. M. C. Schlamp, X. Peng, and A. P. Alivisatos, *J. Appl. Phys.* 82, 5837 (1997).
63. Y. Yang, J. Huang, B. Yang, S. Liu, and J. Shen, *Synth. Met.* 91, 347 (1997).
64. Y. Horii, M. Katagawa, H. Taneoka, H. Kusano, T. Murakami, Y. Hino, and H. Kobayashi, *Mater. Sci. Eng. B* 85, 92 (2001).
65. T. P. Nguyen, P. Le Rendu, M. Lakéhal, P. Joubert, and P. Destruel, *Mater. Sci. Eng. B* 69–70, 177 (2000).
66. S. A. Curran, P. M. Arayan, W. J. Blau, D. L. Carroll, J. N. Colemen, A. B. Dalton, A. P. Davey, A. Drury, B. McCarthy, S. Maier, and A. Strevens, *Adv. Mater.* 10, 1091 (1998).
67. S. Meada and S. P. Armes, *Synth. Met.* 73, 151 (1995).
68. M. Granström, M. Berggren, and O. Inganäs, *Science* 267, 1479 (1995).
69. M. Granström, M. Berggren, and O. Inganäs, *Synth. Met.* 76, 141 (1996).
70. G. E. Jabbour, D. L. Mathine, B. Kippelen, and N. Peyghambarian, *Electron. Lett.* 33, 2070 (1997).
71. K. Mori, Y. Sakaguchi, Y. Iketsu, and J. Suzuki, *Displays* 22, 43 (2001).
72. U. Lemmer, D. Vacar, D. Moses, A. J. Heeger, T. Ohnishi, and T. Noguchi, *Appl. Phys. Lett.* 68, 3007 (1996).
73. G. Yu, Q. Pei, and A. J. Heeger, *Appl. Phys. Lett.* 70, 934 (1997).
74. Y. Kaminorz, E. Smela, O. Inganäs, and L. Brehmer, *Adv. Mater.* 10, 765 (1998).
75. Y. Kaminorz, E. Smela, T. Johansson, L. Brehmer, M. R. Andersson, and O. Inganäs, *Synth. Met.* 113, 103 (2000).
76. D. G. Lidzey, M. A. Pate, M. S. Weaver, T. A. Fisher, and D. D. C. Bradley, *Synth. Met.* 82, 141 (1996).
77. K. Nagayama, T. Yahagi, H. Nakada, T. Tohma, T. Watanabe, K. Yoshida, and S. Miyaguchi, *Jpn. J. Appl. Phys.* 36, L1555 (1997).
78. D. B. Roitman, H. Antoniadis, R. Helbing, F. Pourmizaie, and J. R. Sheats, *Proc. SPIE* 3476, 232 (1998).
79. J. Wang, G. Yu, G. Srdanov, and A. J. Heeger, *Org. Electron.* 1, 33 (2000).
80. C. Hosokawa, M. Eida, M. Matsuura, K. Fukuoka, H. Nakamura, and T. Kusumoto, *Synth. Met.* 91, 3 (1997).
81. P. F. Tian, P. E. Burrows, and S. R. Forrest, *Appl. Phys. Lett.* 71, 3197 (1997).
82. C. Py, D. Roth, I. Lévesque, J. Stapledon, and A. Donat-Bouillud, *Synth. Met.* 122, 225 (2001).
83. P. Visconti, C. Turco, R. Rinaldi, and R. Cingolani, *Microelectron. Eng.* 53, 391 (2000).
84. Y. Xia and G. M. Whitesides, *Angew. Chem. Int. Ed.* 37, 550 (1998).
85. T. Grandlund, T. Nyberg, L. S. Roman, M. Svensson, and O. Inganäs, *Adv. Mater.* 12, 269 (2000).
86. F. Nüesch, Y. Li, and L. J. Rothberg, *Appl. Phys. Lett.* 75, 1799 (1999).
87. F. Nüesch, L. J. Rothberg, E. W. Forsyth, Q. T. Le, and Y. Gao, *Appl. Phys. Lett.* 74, 880 (1999).
88. J. Wang, X. Sun, L. Chen, and S. Y. Chou, *Appl. Phys. Lett.* 75, 2767 (1999).
89. T. Nyberg, F. Zhang, and O. Inganäs, *Nanotechnology* 13, 205 (2002).
90. C. Kim, P. E. Burrows, and S. R. Forrest, *Science* 228, 831 (2000).
91. J. Sone, J. Fujita, Y. Ochiai, S. Manako, S. Matsui, E. Nomura, T. Baba, H. Kawaura, T. Sakamoto, C. D. Chen, Y. Nakamura, and J. S. Tsai, *Nanotechnology* 10, 135 (1999).

92. Z. Bao, J. A. Rogers, and H. E. Katz, *J. Mater. Chem.* 9, 1895 (1999).
93. P. Calvert, *Chem. Mater.* 13, 3299 (2001).
94. R. Hebner, C. C. Wu, D. Marcy, M. H. Lu, and J. C. Sturm, *Appl. Phys. Lett.* 72, 519 (1998).
95. J. Bharathan and Y. Yang, *Appl. Phys. Lett.* 72, 2660 (1998).
96. S. C. Chang, J. Bharathan, Y. Yang, R. Helgeson, F. Wudl, M. B. Ramey, and J. R. Reynolds, *Appl. Phys. Lett.* 73, 2561 (1998).
97. T. Shimoda, S. Kanbe, H. Kobayashi, S. Seki, H. Kiguchi, I. Yudasaka, M. Kimura, S. Miyashita, R. H. Friend, J. H. Burroughes, and C. T. Towns, *Soc. Inform. Display* 379 (1999).
98. P. C. Duinevel, M. M. de Kok, M. Buechel, A. Sempel, K. A. H. Mutsaers, P. van de Weijer, I. G. J. Camps, T. J. M. van den Biggelaar, J. E. J. M. Rubingh, and E. I. Haskal, *Proc. SPIE* 4464, 59 (2001).
99. S. P. Speakman, G. G. Rozenberg, K. J. Clay, W. I. Milne, A. Ille, I. A. Gardner, E. Bresler, and J. H. G. Steinke, *Org. Electron.* 2, 65 (2001).
100. S. Noach, E. Z. Faraggi, G. Cohen, Y. Avny, R. Neumann, D. Davidov, and A. Lewis, *Appl. Phys. Lett.* 69, 3650 (1996).
101. D. A. Pardo, G. E. Jabbour, and N. Peyghambarian, *Adv. Mater.* 12, 1249 (2000).
102. B. Birnstock, J. Bläsing, A. Hunze, M. Scheffel, M. Stössel, K. Heuser, G. Wittmann, J. Wörle, and A. Winnacker, *Appl. Phys. Lett.* 78, 3905 (2001).
103. M. Cölle and T. Tsutsui, *Synth. Met.* 111–112, 95 (2000).
104. R. Micheletto, N. Yoshimatsu, M. Yokokawa, T. An, H. Lee, and S. Okazaki, *Opt. Commun.* 196, 47 (2001).
105. M. Klemenc, F. Meghdadi, S. Voss, and G. Leising, *Synth. Met.* 85, 1243 (1997).
106. H. N. Lin, S. H. Chen, Y. Z. Lee, and S. A. Chen, *J. Vac. Sci. Technol. B* 19, 308 (2001).
107. T. Marks, J. G. C. Veinot, J. Cui, H. Yan, A. Wang, N. L. Edleman, J. Ni, Q. Huang, P. Lee, and N. R. Armstrong, *Synth. Met.* 127, 29 (2002).
108. D. Braun, *Synth. Met.* 92, 107 (1998).
109. D. Braun, J. Rowe, and G. Yu, *Synth. Met.* 102, 920 (1999).
110. Gyoutoku, S. Hara, T. Komatsu, M. Shirinashihama, H. Iwanaga, and K. Sakanoue, *Synth. Met.* 91, 73 (1997).
111. Sempel and M. Büchel, *Org. Electron.* 3, 89 (2002).
112. S. W. Kim, B. H. Hwang, J. H. Lee, J. I. Kang, K. W. Min, and W. Y. Kim, *Curr. Appl. Phys.* 2, 335 (2002).
113. C. Hosokawa, M. Matsuura, M. Eida, K. Fukuoka, H. Tokailin, and T. Kusumoto, *J. SID* 6, 257 (1998).
114. Y. Fukuda, T. Watanabe, T. Wakimoto, S. Miyaguchi, and M. Tsuchida, *Synth. Met.* 111–112, 1 (2000).
115. H. Kubota, S. Miyaguchi, S. Ishizuka, T. Wakimoto, J. Funaki, Y. Fukuda, T. Watanabe, H. Ochi, T. Sakamoto, T. Miyake, M. Tsuchida, I. Ohshita, and T. Tohma, *J. Lumin.* 87–89, 56 (2000).
116. H. Sirringhaus, N. Tessler, and R. H. Friend, *Science* 280, 1741 (1998).
117. A. Dodabalapur, Z. Bao, A. Makhija, J. G. Laquindanum, V. R. Raju, Y. Feng, H. E. Katz, and J. Rogers, *Appl. Phys. Lett.* 73, 142 (1998).
118. R. Ihihara and M. Matsumura, *Jpn. J. Appl. Phys.* 36, 6167 (1997).
119. Z. Meng, C. H. Chen, C. Qiu, H. S. Kwok, and M. Wong, *SID Dig.* 380 (2001).
120. R. M. A. Dawson, Z. Shen, D. A. Furst, S. Connor, J. Hsu, M. G. Kane, R. G. Stewart, A. Ipri, C. N. King, P. J. Green, R. T. Flegal, S. Pearson, W. A. Barrow, E. Dickey, K. Ping, C. W. Tang, S. Van Slyke, F. Chen, J. Shi, J. C. Sturm, and M. H. Lu, *SID* (1998).
121. Y. He, R. Hattori, and J. Kanicki, *Jpn. J. Appl. Phys.* 40, 1199 (2001).
122. H. H. Kim, T. M. Miller, E. H. Westerwick, Y. O. Kim, E. W. Kwock, M. D. Morris, and M. Cerullo, *J. Lightwave Technol.* 12, 609 (1994).
123. C. C. Wu, S. D. Theiss, G. Gu, M. H. Lu, J. C. Sturm, S. Wagner, and S. R. Forrest, *IEEE Electron Device Lett.* 18, 609 (1997).
124. T. Tohma, S. Yamazaki, and D. Wzorek, *Inform. Displays* 11, 20 (2001).
125. R. M. A. Dawson, Z. Shen, D. A. Furst, S. Connor, J. Hsu, M. G. Kane, R. G. Stewart, A. Ipri, C. N. King, P. J. Green, R. T. Flegal, S. Pearson, W. A. Barrow, E. Dickey, K. Ping, S. Robinson, C. W. Tang, S. Van Slyke, F. Chen, J. Shi, M. H. Lu, M. Moskewicz, and J. C. Sturm, *SID* (1999).
126. A. P. Ghosh, *Proc. SPIE* 4464, 1 (2002).
127. J. W. Seo, H. Kim, B. Kim, and Y. Kim, *Jpn. J. Appl. Phys.* 41, 7391 (2002).
128. A. M. Stoneham, *Mater. Sci. Eng. C* 23, 235 (2003).
129. M. A. Ratner, *Mater. Today* 20 (2002).
130. G. D. J. Smit, M. G. Flokstra, S. Rogge, and T. M. Klapwitjk, *Microelectron. Eng.* 64, 429 (2002).
131. M. Iwamoto, *J. Mater. Chem.* 10, 99 (2000).
132. Y. Aoyagi and T. Sugano, *Supramol. Sci.* 3, 177 (1996).
133. A. J. Epstein in "Organic Electronic Materials" (R. Farchioni and G. Grosso, Eds.), p. 3. Springer, Berlin, 2001.
134. C. Xu, Q. Xue, Y. Zhong, Y. Cui, L. Ba, B. Zhao, and N. Gu, *Nanotechnology* 13, 47 (2002).
135. J. M. Lupton, I. D. W. Samuel, R. Beavington, P. L. Burn, and H. Bässler, *Synth. Met.* 116, 357 (2001).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanostructured Surface Modifications for Biomedical Implants

Shane A. Catledge, Marc Fries, Yogesh K. Vohra

University of Alabama at Birmingham, Birmingham, Alabama, USA

CONTENTS

1. INTRODUCTION

With the advancement of modern materials science, many times a single material does not possess the desired properties for certain applications. For example, some materials have desired bulk properties (chemical, mechanical, electrical, etc.) but their surface properties are inadequate. In these cases, a combination of two or more materials or a modification of the bulk material is adopted. Surface modification is basically a process in which the surface of a material is chemically, physically, or mechanically altered so that the surface properties are changed from that of the original bulk. Materials surfaces are commonly modified by techniques involving, for example, deposition of some additional material as a thin coating or implantation of energetic ions to alter the local atomic bonding. In addition, multilayer or functionally graded structures can be engineered to provide a more viable transition from the bulk to the surface.

Novel materials whose microstructure can be engineered to contain features with nanometer-scale dimensions have become quite popular, especially in the last two decades. These "nanostructured" materials can exhibit enhanced mechanical, electrical, magnetic, and/or optical properties compared to their conventional micrometer-scale (or larger) counterparts. Nanostructured (NS) materials contain a large volume fraction (greater than 50%) of defects such as grain boundaries, interphase boundaries, and dislocations, and this strongly influences their chemical and physical properties. The Vickers hardness of electrodeposited nickel, for example, increases from 140 to 650 when the grain size is reduced from the microcrystalline range to about 10 nm [1]. Similarly, NS ceramics tend to be tougher and stronger than coarser grained ceramics. The synthesis and control of materials in nanometer dimensions can provide access to new levels of material properties and device characteristics that were previously unattainable, and work is rapidly expanding worldwide in attempts to exploit the opportunities offered through nanostructuring.

Terms with the prefix "nano"—nanostructured, nanocrystalline, and nanophase, for example—have often been used interchangeably to describe materials with nanometer-scale dimensions. However, it is important to distinguish between different structural forms of materials containing nanometer-size features. Defined broadly, the term *nanostructured* is reserved for materials characterized by structural features of less than 100 nm in average size. Nanostructured materials can take the form of powders, dispersions, coatings, or bulk materials. One rapidly evolving class of NS materials is that of nanocomposites in which nanometer-size particles or whiskers act as reinforcements in a matrix composed of conventional-size grains or amorphous material. For example, hard composite coatings composed of nanocrystalline (10–50 nm) TiC grains imbedded in an amorphous carbon matrix (about 30 vol%) have shown a fourfold increase in toughness compared to nanocrystalline single-phase TiC [2]. Such coatings can be designed to provide adequate compliance through the controlled formation of dislocations and nano- and microcracks once stresses exceed the elastic limit, while still maintaining high hardness and eliminating catastrophic failure typical of hard, brittle materials. For purposes of this chapter and to avoid confusion with other "nano" terminology, we define *nanocrystalline* as a more specific subcategory of NS materials characterized by nanometer-size grains without the presence of conventional-size grains or an intergranular amorphous phase.

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (741–762)

The tremendous impact on the biomedical community by the unique properties offered by NS materials has stimulated intense research to evaluate the possibility of using them in the design of orthopedic and dental implants.

1.1. Orthopedic Implants

Over 450,000 total knee and total hip replacement operations were performed in the United States in 1994 with an anticipated increase of 10% per year [3]. The typical useful life for a replaced joint is between 10 and 15 years [4]. Complications arising from wear include component loosening, deleterious biological responses, osteolysis, mechanical instability, decreased joint mobility, increased pain, and ultimately implant failure. Therefore, revision surgery is frequently necessary, particularly in younger, more active patients. If the useful life of the implant were extended (to upward of 40 years) considerable patient suffering could be eliminated and a substantial amount of health care dollars could be saved. It has been shown [5] that the wear rate of implant components is noticeably lower when articulated against ceramic materials such as zirconia than for metallic materials. Unfortunately, zirconia devices are three times more expensive than their metallic counterparts. Common drawbacks to ceramic devices are brittleness, geometrical limitations, and catastrophic failure modes. These physical limitations have made manufacture of hips difficult and knees impossible; there are currently no available femoral knee components made from solid ceramics. Thus, the major goal is to develop coating technology that can reduce the friction and wear in mating total joint replacement components, thus contributing to their significantly improved function and longer life span.

1.2. Dental Implants

It is estimated that currently there are approximately 10 million people in America, the majority of them women, who suffer pain and dysfunction in and near the temporomandibular joint (TMJ). In previous decades, TMJ implants were fabricated of polymeric materials, including the Vitek and Silastic TMJ implants. The primary problem area with these implant systems was wear and deterioration of the components. The almost 100% failure rate of these implants after 3 years lead to a recall of these devices. Revision surgery is needed in most of these patients due to excessive loss of bone and function, and there is currently no fully successful implant system available. Currently, the two devices most commonly used for TMJ restoration are: (1) cobalt–chrome metal on metal devices and (2) polyethylene condyle on a titanium alloy shaft articulating against a curved cobalt–chromium plate (glendoid-fossa component) as shown in Figure 1. Although the patient's own bone can be used to rebuild a portion of the deficient area around the TMJ, lost function is not restored in many cases and an artificial device is necessary. Metal-to-metal TMJ implants fabricated to date have also exhibited loosening problems because of the type of sliding and rotational motions involved, and metal debris is seen in fairly high amounts in the joint area. Another problem with the TMJ devices which leads to higher than expected wear (primarily the fretting type) is the lack of stable fixation in bone, as retaining screws can come loose allowing the implant move out of its ideal position. The ceramic coatings discussed in this chapter will address the issues of wear and loosening in these devices by nanostructural modification of the implant surfaces, which will allow improvement of existing articulations utilizing metal–metal and metal–polymer implants. This approach will enhance screw and component fixation through bone-to-hydroxylapatite interactions and hopefully lead to the development of a more successful TMJ implant design. Further background can be obtained from reviews on the subject (see [6, 7]).

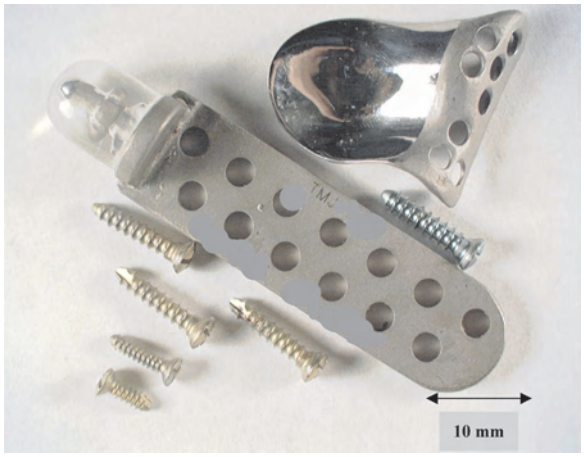


Figure 1. The components of a temporomandibular joint dental implant. The photograph shows a long titanium alloy shaft with a round condylar head made out of polyethylene. This polyethylene condyle articulates against a cobalt–chromium curved metal plate (glendoid-fossa component) shown in the top right hand corner of the photograph. The Ti-alloy screws shown are used to attach the shaft to the supporting bone. We have selected this device as a case study for nanostructured bioengineered diamond and calcium phosphate ceramic coatings. Reprinted with permission from [148], S. A. Catledge et al., *J. Nanosci. Nanotechnol.* 2, 293 (2002). © 2002, American Scientific Publishers.

In this chapter, we focus on the following areas of cutting edge research:

1. Nanostructured bioengineered diamond coatings on cobalt–chrome and titanium alloys that will increase the life span of a variety of implant devices to upward of 40 years. Nanostructured diamond and diamondlike carbon coatings have great potential for use as biomedical implants due to their extreme hardness, wear resistance, low friction, and biocompatibility characteristics. Nanostructured diamond produced by chemical vapor deposition (CVD) techniques and comprised of nanosize diamond grains has particular promise because of the combination of ultrahigh hardness, improved toughness, and good adhesion to titanium alloys. The structure and properties of these coatings are easily tailored to the desired application by appropriate choice of feedgas chemistry and other deposition conditions.
2. Nanostructured hydroxyapatite coatings on metals that will promote better attachment to the bone while optimizing abrasion resistance, bond strength, and the dissolution rate. Hydroxyapatite (HA) is a bioactive

ceramic with a crystal structure similar to native bone and teeth minerals. It has generated great interest in the search for advanced orthopedic and dental implant materials as it elicits a favorable biological response and forms a bond with the surrounding tissues. However, applications of HA are currently limited to powders, coatings, porous bodies, and non-load-bearing implants due to processing difficulties and the poor mechanical properties of conventional HA. Nanostructured processing applied to hydroxyapatite-based materials is used to achieve the desired mechanical characteristics and enhanced surface reactivity for multifunctional implant systems tailored toward different hard tissue replacements. It has allowed chemical homogeneity and microstructural uniformity to be achieved for HA so that fully dense bioceramics can be generated at low sintering temperatures with a significant reduction in flaw size. The nanometer-sized grains and high volume fraction of grain boundaries in nanostructured HA have been found to increase osteoblast adhesion, proliferation, and mineralization.

3. Nanostructured metalloceramic coatings that will provide continuous variation from a nanocrystalline metallic bond at the interface to the hard ceramic bond on the surface. Bonding of functionally graded, NS metalloceramic coatings gradually changes from metallic to primarily covalent with increasing ceramic material toward the surface. One advantage of the graded bonding structure is the ability to overcome adhesion problems associated with ceramic hard coatings on metallic substrates while exhibiting enhanced surface hardness and wear resistance. The NS metalloceramic coatings have a smooth surface finish that is retained *in vivo*, leading to a higher wear resistance.

2. NANOSTRUCTURED DIAMOND COATINGS

Diamond growth from the gas phase was reported as early as 1911 by von Bolton [8] using acetylene in the presence of mercury vapor at 100 °C. However, the first documented report of diamond growth at low pressures, which was subsequently confirmed by others, was the work of William G. Eversole at the Union Carbide Corporation. In the period between 26 November 1952 and 7 January 1953, Eversole achieved diamond growth on a diamond seed crystal using carbon monoxide as a source gas. These findings were documented in 1956 [9]. This predates the first documented, reproducible synthesis of diamond by a high pressure, high temperature process by scientists at the General Electric Company in 1955 [10]. The enormous success of this technique of producing diamond where it is thermodynamically stable, along with the impractically low growth rates initially observed for low pressure synthesis of diamond, delayed further development of the low pressure techniques. There also still remained a strong feeling that diamond grown at pressures where it is a metastable phase was somehow thermodynamically forbidden. It was not until the late 1970s, when advances were made in the USSR [11, 12] for yielding continuous films using hydrogen and methane mixtures and in Japan [13] for rendering economically viable growth rates, that the scientific community was finally convinced of the feasibility of low pressure diamond synthesis.

It is well known that diamond is the hardest known material. In fact, it is in the market of abrasives and cutting tools that diamond has found its most prominent and economically sustaining niche. Other conventional, hard bulk materials such as cubic boron nitride can reach, at best, about 47 GPa (only about 47% diamond hardness). Granted, appropriate engineering of the chemistry and microstructure of materials has led to a new class of synthetic superhard materials with exciting potential. For example, epitaxial [14] and polycrystalline [15] superlattices of transition metal nitrides can reach hardness of 50 GPa when the lattice period decreases to 5–7 nm. Carefully designed composites containing nanocrystals of a hard transition metal nitride imbedded into an amorphous silicon nitride matrix has led to claims of achieved hardness greater than 50 GPa with the possibility of approaching the 70–100 GPa range for diamond [16, 17]. However, the extremely wide range of synthetic carbon structures attainable by low pressure deposition techniques has resulted in coatings with measured hardness ranging from 10 (amorphous carbon) to 100 GPa (crystalline diamond) [18–20]. For this reason, carbon-based coatings with a myriad of different structural forms and ascribed names such as diamond, diamondlike carbon, nanostructured diamond, and tetrahedral amorphous carbon continue to dominate the superhard material arena. It is not surprising that these coatings are particularly suited for, and to some extent already commercially realized, in wear resistant applications such as in the machining of nonferrous materials using diamond-coated carbide cutting tools. Recent advances have led to NS diamond coatings exhibiting a unique combination of high hardness, high toughness, and low friction. These and other carbon coatings continue to be actively investigated with respect to the scientific issues (growth mechanisms, structure, etc.) and technological issues (large-area deposition, adhesion onto metallic substrates, etc.) involved. While there are several deposition techniques for producing hard carbon films (e.g. ion beam, pulsed laser, cathodic arc, etc.), we will focus on the use of CVD as it pertains to NS diamond growth.

Chemical vapor deposition using hydrogen-rich, hydrocarbon-containing gases has been the most successful method of producing metastable diamond films of high crystallinity. Work by Lander and Morrison [21] at Bell Labs showed in 1966 that a diamond surface saturated with hydrogen maintains the bulk terminated diamond lattice to the outermost surface layer of carbon atoms. Without the hydrogen atoms to terminate the carbon dangling bonds, the surface reconstructs into more complex structures. In 1971 Angus et al. [22] showed that atomic hydrogen could be used as a cyclic etching step to remove unwanted graphite deposited during CVD of diamond. However, the importance of atomic hydrogen used during the diamond growth process itself was realized predominantly by Fedoseev et al. [23] in the USSR, during the mid to late 1970s. In this way, atomic hydrogen could simultaneously (i) stabilize the diamond growth surface, (ii) preferentially etch graphitic and other nondiamond carbon, and (iii) create new growth sites on the surface via hydrogen-abstraction reactions.

This allowed much higher growth rates and permitted the nucleation of new diamond crystallites on nondiamond substrates. The reader is referred to a review by Angus and Hayman [24] for a more comprehensive description of the diamond growth process involving hydrogen-rich, hydrocarbon-containing gases and the corresponding roles of atomic hydrogen.

2.1. Structure

A conventional mixture used to produce high phase purity microcrystalline diamond films by CVD using hydrocarbon-containing precursors is about 1 to 3% methane in a balance of hydrogen. Other deposition parameters such as substrate and gas-phase temperature, as well as the choice of substrate material and how it its surface is prepared, can affect the film structure and morphology. However, the former mixture will typically produce a rough, faceted diamond surface comprised of micrometer-size grains. For many applications, a rough surface is undesirable as it can cause severe abrasion during sliding against other materials. By appropriate choice of feedgas types and concentrations, it is possible to manipulate the diamond growth process to produce NS films with high hardness and smooth, wear-resistant surfaces. Figure 2 shows an example of the difference in microstructure between microcrystalline and NS diamond films produced by microwave plasma CVD. Both films were grown by Catledge and Vohra [25] on a titanium alloy (Ti-6Al-4V) using an unconventionally high methane concentration of 15%. The microcrystalline film was grown under plasma conditions allowing sufficient production of atomic hydrogen via elevated gas pressure (125 Torr), along with very low levels of nitrogen impurity in the feedgas. In this way, the high methane concentration can lead to an increase of nucleation and growth rates without a substantial decrease in diamond quality. An added benefit of the high methane concentration is an improvement of diamond-to-titanium adhesion when compared to the more conventional 1 to 3% mixture [25, 26]. Adhesion of conventionally grown diamond films on metallic substrates is a concern because large thermally induced residual stresses provide a critical driving force for the propagation of cracks in the rather unforgiving and brittle nature of the film. In this example, the nanostructured nature of the film shown in the right side of Figure 2 is induced by the addition of large amounts (10% of methane flow) of nitrogen to the feedgas mixture. The grain size of the NS film as measured by X-ray diffraction is 15 nm and the surface roughness is 27 nm root-mean-squared (rms) value, an order of magnitude less than the microcrystalline film. By improving the surface finish of the initial predeposited substrate to within 10 nm (rms value), we have achieved NS diamond films with surface roughness as low as 15 nm [26]. As compared to the microcrystalline film, the Raman spectrum of the NS film reveals a weaker diamond peak (ca. 1332 cm^{-1}), pronounced scattering in the 1350–1550 cm^{-1} region associated with an increase of sp^2-bonded carbon, and an additional peak at 1140 cm^{-1}. It should be noted that the intensity scales on the Raman spectra of Figure 2 are not the same and that the diamond peak intensity for the NS film is a factor of three lower than that for the microcrystalline film. It has been shown that the weaker diamond peak is caused by a reduction of grain size to the nanometer scale [27], and that the increased sp^2-bonded carbon is due to increasing π-bonds at the grain boundaries in the nanocrystalline films [28]. The additional peak at 1140 cm^{-1} has been attributed to scattering from nm-sized diamond grains [29].

As measured by X-ray photoelectron spectroscopy (XPS), the NS diamond films grown by adding large amounts of nitrogen to the feedgas mixture do not contain a substantial nitrogen concentration. The elemental depth profile by XPS is shown in Figure 3 for a NS diamond film grown on Ti-6Al-4V alloy. Each scan step included oxygen 1*s*, vanadium 2*p*, titanium 2*p*, nitrogen 1*s*, bulk carbon, carbidic carbon, and aluminum 2*p* specific scans. The collected data

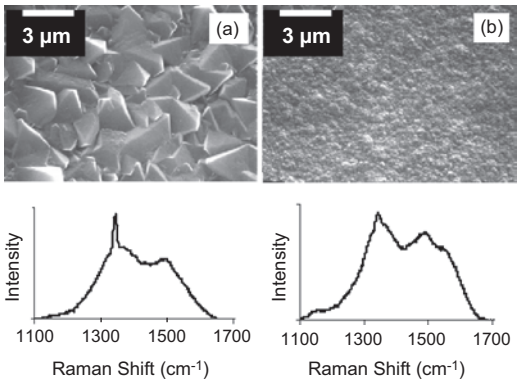


Figure 2. Scanning electron microscope images and Raman spectra of CVD diamond films grown with mixtures of (a) H_2/CH_4 with flow ratio of 500/88 and (b) $H_2/CH_4/N_2$ with flow ratio of 500/88/8.8. Both films were grown using a chamber pressure of 125 Torr and an unconventionally high CH_4 fraction (15 vol%). The smooth nanostructured film in (b) results from the addition of nitrogen to the feedgas mixture. Reprinted with permission from [148], S. A. Catledge et al., *J. Nanosci. Nanotechnol.* 2, 293 (2002). © 2002, American Scientific Publishers.

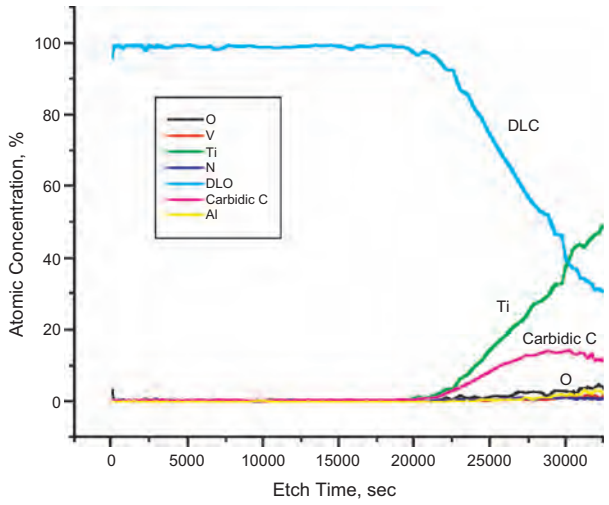


Figure 3. Elemental concentration depth profile of NS diamond film measured by XPS. The three predominant elements in the interface region on the right side of the graph include film carbon (labeled "DLC" here), Ti, and carbidic carbon. The profile represents a diffuse film/substrate boundary.

show the classic pattern for a diffuse interlayer boundary, with bulk carbon decreasing in proportion to TiC and metallic titanium with increasing depth. The proportion of carbidic carbon in the sample decreases starting at the point where bulk titanium overtakes bulk carbon as the primary species detected. The ratio of titanium to carbidic carbon is roughly 2:1 from the onset of detection of titanium up to the point where it makes up about 30% of the total volume, indicating that metallic titanium is present along with TiC throughout the interface region. Oxygen is present in the interface region at a mean value of 3% for the last 8160 s of sputtering time. This finding indicates that the titanium surface is not completely etched bare during exposure to the hydrogen plasma. It is likely that amorphous, nonstoichiometric titanium oxides exist within the interface as has been detected in other diamond/titanium systems. The XPS data reveal only trace quantities of nitrogen, suggesting that the dramatic changes observed in film morphology and structure with nitrogen feedgas addition are due not to nitrogen incorporation, but to the result of gas-phase and surface reactions involving nitrogen related species such as CN.

Due to the small grain size of NS diamond films, a significant fraction of the carbon atoms is situated in the grain boundaries. It is believed that the grain boundary carbon is responsible for the absorption and scattering of light in these films, their electrical conductivity, and their electron emission properties [30]. Simulations by Keblinski et al. [31] have revealed important structural differences between the grain boundaries of phase pure nanocrystalline diamond and those of silicon, and these findings may partially explain the remarkable properties of nanocrystalline diamond films. In contrast to silicon, carbon can form both sp^2- and sp^3-hybridized electronic states. The greater bond stiffness of diamond combined with its ability to change hybridization in a disordered environment results in more ordered grain boundary structures in diamond than in silicon. This comes at the expense of many carbon atoms in the grain boundary (typically 70–80%) being only threefold-coordinated. The competition between structural disordering and local hybridization change, which is absent in silicon, implies that the high degree of structural disorder in the mostly fourfold-coordinated silicon grain boundaries is replaced by bond-coordination disorder in the mostly threefold-coordinated, structurally less disordered diamond grain boundaries. Due to the different numbers of bonds per unit area across the grain boundary and degrees of sp^2 hybridization on high and low energy grain boundary planes, it was shown that nanocrystalline diamond (likely to contain a large fraction of high energy grain boundaries) may be mechanically stronger than coarser grained diamond (containing a significant fraction of low energy grain boundaries) or even perfect crystal diamond. The implication of this is that the dominating brittle-fracture mode in nanocrystalline diamond might be intra- rather than intergranular. Finally, Keblinski et al. showed that the threefold-coordinated grain boundary carbon atoms were poorly connected to each other, and therefore any type of graphitelike electrical conduction through the boundaries is not likely unless "bridging" impurities provide the missing connections.

A final note regarding structure involves a discussion of how the nm-sized diamond grains may be distributed in the film and whether a second or third phase is present beyond the grain boundaries. In other words, it may be possible to have a nanocomposite film comprised of nm-sized diamond (and perhaps also graphitic) grains imbedded in a surrounding amorphous carbon matrix. Such films, if designed properly, could be remarkably hard and tough. To the authors' knowledge, the only reported nanocomposite-like films based purely on carbon have been achieved through postdeposition annealing of diamondlike amorphous carbon (*a*-C) films [32]. This results in nm-sized regions of high density *a*-C embedded in a lower density *a*-C matrix. The hardness in this case was reported to increase by 15% (to 88 GPa) due to the nanocomposite structure. A few other reports have shown diamond nanocrystallites sporadically imbedded within amorphous carbon [33, 34]. However, a uniform microstructure with controlled dispersion of a nanodiamond phase within a surrounding amorphous carbon matrix is not yet reported.

2.2. Processing Routes

A range of processing routes to CVD NS diamond films exist, but they all have in common the deposition of nanometer-size grains and a film surface roughness that is typically about an order of magnitude lower than that of microcrystalline diamond films. It is possible that early observations of nanocrystallinity, or at least of the apparent degradation of diamond film quality resulting from nanocrystallinity, was concluded to be a result of gas-phase contamination by nitrogen. It is well known that nitrogen is a common impurity in natural diamond and that it can influence physical properties such as optical transparency as well as electrical and thermal conductivity [35]. In 1993, Badzian et al. [36] reported deliberate use of N_2 in place of H_2 to determine its effect on CVD diamond growth. They concluded that the mixture of $CH_4 + N_2$ resulted in a disordered diamond structure likely to be caused by surface processes involving atomic nitrogen. Nitrogen was not incorporated into the film to the amount detectable by Auger electron spectroscopy and XPS. Although not identified as containing nanocrystalline diamond grains, Raman spectroscopy of the films produced by Badzian et al. have similar features to those of NS diamond. In addition, their diamond X-ray diffraction (XRD) peaks were broad, suggestive of reduced grain size. Work by Locher et al. [37] and Jin and Moustakas [38] followed to study the influence of varying amounts of nitrogen in conventional $CH_4 + H_2$ mixtures for CVD diamond growth and revealed dramatic changes in structure and morphology. The addition of small amounts (40–200 ppm) of nitrogen results in a pronounced $\langle 100 \rangle$ texture with nearly coplanar surface facets and improved crystalline diamond quality as determined by Raman spectroscopy. It was established that much higher nitrogen additions led to a deterioration of film quality which according to Locher et al., became "nanocrystalline." The potentially beneficial properties of such films (such as improved wear resistance and electron emission) appear not to have been recognized at that time. In fact, the films may have been determined to be inferior due to the apparent lack of diamond phase purity.

Since the work of Locher et al., many researchers have investigated the influence of nitrogen addition on CVD diamond growth and realized that unique properties (such as enhanced electron field emission [39] and improved toughness [26]) can result from the nanostructured films. The evolution from micro- to nanocrystalline diamond grains has prompted considerable investigation into the mechanisms responsible for the nitrogen-induced transformations. For hydrogen-rich reactant gases, the importance of methyl radicals to the diamond growth process is well known [40]. It is likely that the relative concentration of CH_3 and CN species is of crucial importance to the growth mechanism and observed changes in morphology. According to Bohr et al. [41], for small N_2 additions CN and HCN can efficiently abstract adsorbed H atoms creating vacant growth sites and improving diamond crystal quality by reducing carbon supersaturation. With larger N_2 additions, the increased supply of CN and HCN enhances the abstraction of adsorbed H leading to reconstruction of the diamond surface, because adsorbed CN or nitrogen species are not able to stabilize the diamond structure at the surface efficiently. This "reconstruction" appears to manifest itself as a reduction of crystal size with a corresponding increase of sp^2-bonded carbon, most of which is presumably located at the grain boundaries. Recent gas-phase simulations [42] indicate that there exists a critical N_2/CH_4 feedgas concentration, above which changes in the CH_3/CN ratio are minimal. The trend in the modeled CH_3/CN ratio corresponds well to the experimentally observed changes in diamond crystallinity and surface roughness [25]. This suggests that above a critical N_2/CH_4 ratio, saturation of CN will not induce any further change in the surface structure. While it has been shown that nitrogen incorporation is higher in {111} growth sectors than in {100} sectors by a factor of 3–4 leading to different growth rates of these faces [43], the overall low level of nitrogen incorporation ($<9 \times 10^{20}$ cm^{-3}) into the diamond [36, 38] suggests that the roles nitrogen play on the growth surface are responsible for the observed changes in structure and growth rates. Possible nitrogen-related surface processes include an increase in the generation of growth steps [44] (especially on {100} faces} or multiple twinning [45] of diamond crystals.

A unique processing route to diamond film growth by CVD was reported by Gruen et al. [46] in 1994 using argon-rich plasmas with fullerenes, but without hydrogen addition. These films were initially characterized as containing phase-pure, fine-grained (0.05–0.3 μm grain size) diamond without graphitic or amorphous phases. Nanocrystalline films were also grown using $Ar/H_2/C_{60}$ [47, 48], $Ar/H_2/CH_4$ [49, 50], Ar/CH_4 [51], N_2/CH_4 [39], and $Ar/CH_4/N_2$ [52] plasmas resulting in grain sizes from 3 to 20 nm and surface roughness as low as 18 nm. Figure 4 shows a plan-view transmission electron microscope (TEM) image of a typical nanocrystalline diamond film produced using 1% CH_4 + 99% Ar. According to Gruen et al., the grain size of this film ranges from 3 to 20 nm. It was observed for argon-rich (>90% Ar) plasmas and elevated reactant gas pressures (100 Torr) that the C_2 concentration is extremely intense compared to atomic hydrogen, while the concentration of methyl radicals is immeasurably low [53]. This is in contrast to the case of conventional (hydrogen-rich, low methane) plasmas

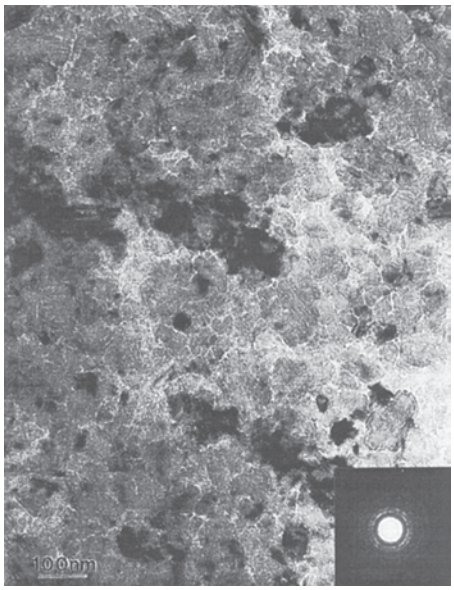


Figure 4. Plan-view TEM image of a diamond film prepared from an Ar/CH_4 plasma at 100 Torr showing that the diamond film consists of nanocrystalline grains ranging from 3 to 20 nm. The inset image shows a sharp ring pattern of selected area electron diffraction, indicating that the diamond grains have a random orientation. Reprinted with permission from [51], D. Zhou et al., *J. Appl. Phys.* 83, 540 (1998). © 1998, American Institute of Physics.

in which the CH_3 radical predominates. The relatively high C_2 plasma emission compared to atomic hydrogen or CH species is illustrated in an optical emission spectrum (Fig. 5) for an Ar-rich plasma (99% Ar, 1% CH_4) at a gas pressure of 100 Torr. The film growth rate and C_2 emission intensity were also observed to increase with increasing reactant gas pressure. These results, along with the fact that the films can be grown in hydrogen-poor or hydrogen-free plasmas, led to a modified explanation for growth involving the carbon dimer C_2, which is believed to be the precursor for nucleation and growth of nanocrystalline diamond in this composition regime. It was suggested that with C_2 as the growth species, no hydrogen abstraction reactions are required, partly because insertion of C_2 into surface C–C bonds is energetically favorable with low activation barriers.

Growth rates for the nanocrystalline diamond films produced from argon-rich plasmas tend to be low, especially for compositions considered to be optimum for yielding small grain size and low surface roughness. For example, conditions required to reduce the rms surface roughness of the film to about 30 nm (e.g., 97% Ar, 2% H_2, 1% CH_4) yield a growth rate of only about 0.3 μm/hr [50]. According to Gruen, the reduction in growth rate may be due in part to the activation energy associated with forming a nucleus of critical size [54]. In order to produce NS diamond films at significantly higher growth rates, and to achieve better adhesion and toughness to metal-carbide forming substrates, Catledge and Vohra developed a process [55] using a mixture of unconventionally high CH_4 concentration (15% of total flow) with N_2 (1.5% of total flow) in a balance of H_2 at a reactant gas pressure of 125 Torr and substrate temperature of 800–850 °C. In this way, NS diamond films with rms roughness of as low as 15 nm can be grown [26]. The C_2

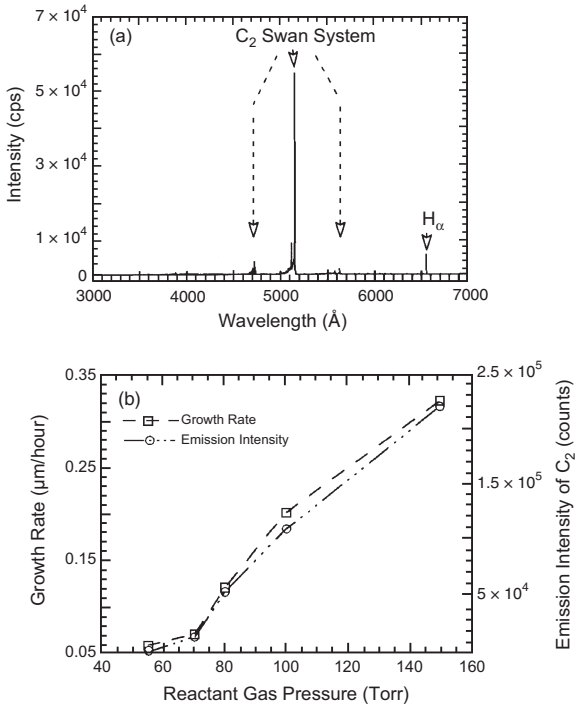


Figure 5. An optical emission spectrum of an Ar/CH_4 plasma running with 1 vol% CH_4 and 99 vol% Ar as the reactant gas at a chamber pressure of 100 Torr; (b) plots of growth rates of nanocrystalline diamond films and emission intensities of C_2 from the Ar/CH_4 plasmas vs reactant gas pressures ranging from 55 to 150 Torr. Reprinted with permission from [51], D. Zhou et al., *J. Appl. Phys.* 83, 540 (1998). © 1998, American Institute of Physics.

emission was found to increase with increasing CH_4 flow and with increasing gas pressure [56], leading to higher nucleation and growth rates. The growth rates for micro- and nanostructured films grown with the 15% CH_4 concentration were 2.4 and 1.8 μm/hr, respectively [57]. The strong influence of the CN radical in causing nanocrystallinity was confirmed by the correlation of its modeled composition in the gas phase [42] with the degree of nanocrystallinity as determined experimentally [18] for diamond films grown with different N_2 additions. For a given CH_4 feedgas concentration, there exists a critical N_2 feedgas concentration, above which the change in the CH_3/CN ratio is minimal. The change in the CH_3/CN ratio with increasing N_2/CH_4 ratio is shown in Figure 6 and illustrates a "knee" in the curves, defining the critical N_2 concentration. The knee becomes more evident when using a higher CH_4/H_2 feedgas ratio. It was observed experimentally that the same critical N_2 feedgas concentration exists, above which a further decrease in diamond crystallinity and surface roughness of the grown diamond films is minimal.

Recently, Sharda et al. [58] have produced NS diamond films with surface roughness in the 15–30 nm range by using biased enhanced growth (BEG) in a microwave plasma CVD system. This unique technique relies on a combination of two growth processes acting simultaneously during deposition. One process involves the surface-related hydrogen abstraction reactions typical for diamond deposition using conventional hydrogen-rich plasmas. In addition, shallow

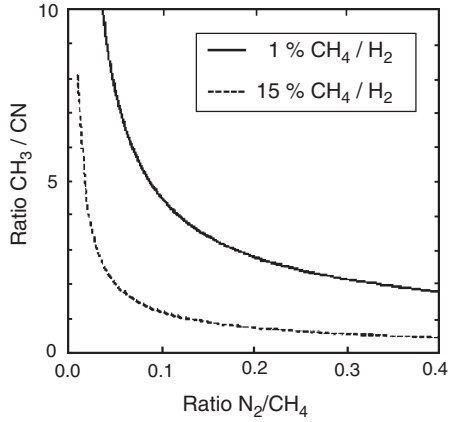


Figure 6. Gas-phase equilibrium plot of the ratio of CH_3 to CN as a function of N_2/CH_4 feedgas ratio as simulated for conditions of diamond growth using an $H_2/CH_4/N_2$ mixture with 1 and 15 vol% CH_4. The chamber pressure was 125 Torr. Reprinted with permission from [42], R. Corvin et al., *Appl. Phys. Lett.* 80, 2550 (2002). © 2002, American Institute of Physics.

subsurface implantation of energetic carbon ions caused by the continuous application of a dc bias voltage to the substrate produces a disordered structure containing a high concentration of sp^3 carbon in the films. The BEG process takes advantage of subsurface implantation of carbon ions to produce a smooth NS diamond film while the process involving hydrogen abstraction reactions helps to etch away graphitic carbon deposits. Relatively low substrate temperatures (400–700 °C) were shown to be useful for this growth regime in obtaining optimal NS diamond content and film hardness.

While the three major processing routes for NS diamond growth discussed in this chapter can be used to produce hard and smooth films, there are likely to be differences in structure, primarily with regard to the relative amounts of sp^3- and sp^2-bonded carbon, the degree of crystallinity, and the amount of hydrogen incorporation in the film. These factors will no doubt influence the mechanical properties of the film such as hardness, adhesion, and wear resistance.

2.3. Mechanical Properties

As grain size decreases to the nm range, the mechanisms for plastic deformation change. In particular, dislocation activity is reduced and the creation of new dislocations is made difficult as the grain size reaches the lower end of the nanoscale (<10 nm). Stresses needed to activate dislocation sources, such as the Frank–Reed source, are inversely proportional to the distance between dislocation pinning points. Since nanoscale grains will limit the distance between such pinning points, the stresses needed to activate dislocation sources can reach the theoretical shear stress of a dislocation-free crystal at the smallest grain sizes (~2 nm). Therefore, current deformation models cannot simply be scaled, but instead new models have to be developed which take the special microstructural features of NS materials into account. It has been suggested that the new deformation mechanisms acting for nanoscale materials may involve grain boundary sliding and formation of mesoscopic glide planes [59].

According to the empirically based Hall–Petch effect, which is due to dislocation motion/generation in materials that exhibit plastic deformation, the strength and hardness of polycrystalline materials increases with decreasing crystallite size. It is well established that as grain size is reduced through the nanoscale regime (<100 nm), hardness typically increases and can be factors of 2 to 7 times higher for pure nanocrystalline metals (10 nm grain size) than for large-grained (>1 μm) metals [60]. For typical ceramics, of course, strength is particularly influenced by Griffith "flaws" such as pores or other heterogeneities common to the manufacturing process. The so-called inverse Hall–Petch effect has been observed for some nanocrystalline materials and results in a decrease in hardness with decreasing grain size below about 10 nm. The transition from the conventional to inverse Hall–Petch effect may be attributed to a competition between dislocation dominated processes and grain boundary sliding controlled processes [59]. These are two independent mechanisms and the one that requires less stress (energy) will be the favored mode of deformation. The softening of some materials below a grain size of about 10 nm may also logically be explained by the strong increase in the amount of disordered grain boundary material, whose strength is lower [61].

The importance of NS composites in improving not only strength and hardness but also fracture toughness has become increasingly recognized. Veprek et al. [17] have developed new superhard nanocrystalline coatings consisting of about 4 nm or less nanocrystals of a hard transition metal nitride embedded into <1 nm thin matrix of amorphous silicon nitride. Unlike pure nanocrystalline materials which show softening when the crystallite size decreases below about 5 nm, the hardness of these nanocomposites strongly increases with decreasing crystallite size in that range and approaches the hardness of diamond. The high hardness observed for many nanocomposite materials appears to be due to the harder phase acting as an obstacle to dislocation movement and grain boundary sliding. This, in turn, is influenced by the number and dispersion of the harder phase. Other nanocomposite coatings such as those based on TiC crystals of 10–50 nm in size embedded into an amorphous carbon phase have shown hardness of 32 GPa and a fourfold increase in toughness in comparison with nanocrystalline single-phase TiC [2]. Clearly, the high hardness/strength and toughness/ductility that appear to be characteristic of recently studied multiphase ceramic nanocomposites are leading to very exciting engineering of material structures.

With regard to diamond films and measurement of their mechanical properties, a critical issue involves how to obtain accurate, reproducible data that are not influenced by the deformation (elastic or plastic) of the testing apparatus itself. However, if one is careful to monitor changes in the surfaces involved, useful data can be obtained. The large variation in film structure that can occur using different processing routes for NS diamond growth requires measurements of film hardness, wear resistance, adhesion, fracture toughness, and friction coefficient.

Catledge et al. [18] have recently studied changes in NS diamond structure and hardness for films produced on Ti-6Al-4V using different N_2/CH_4 feedgas ratios. The relative concentration of CH_4 and N_2 (in a balance of H_2) is shown to strongly influence film structure, hardness, and adhesion. For high CH_4 concentration (15% by volume), NS diamond films with roughness of 20–30 nm, good adhesion, and high hardness (85 GPa) are obtained. As shown in Figure 7, a distinct correlation was found between the nanoindentation hardness of the deposited film and the N_2/CH_4 ratio in the plasma as well as correlation of hardness to the ratio of the Raman peak intensities (1332 and 1555 cm^{-1}). The most abrupt changes in film structure occur in the low N_2/CH_4 ratio range less than 0.10. Above this ratio, the crystalline diamond component does not change significantly and the surface roughness and friction coefficient are not lowered any further. This may be explained by observations that the concentration of CN radical in the plasma saturates above a N_2/CH_4 feedgas ratio of 0.10.

The measured hardness of the biased enhanced growth of NS diamond films produced by Sharda et al. [58] is shown in Figure 8 as a function of substrate temperature. Also plotted with temperature is the Raman intensity ratio of nanocrystalline diamond (I_n @ 1150 cm^{-1}) and graphitic (I_g @ 1580 cm^{-1}) components. Substrate temperature is found to play a crucial role in the growth of nanocrystalline diamond by this technique and the maximum hardness (~40 GPa) is reached at about 600 °C. It is suggested that higher temperatures cause the implanted carbon atoms to migrate back to the surface resulting in domination of surface processes in the growth, which in turn results in an increase in the graphitic content of the film. By optimizing the growth parameters in this technique, film hardness can be increased to above 75 GPa.

The adhesion of diamond films to the substrate is a very important issue, especially for deposition on metallic substrates where thermally induced residual film stress can provide a large driving force for delamination. Several techniques have been used to evaluate adhesion and fracture toughness of diamond films [62–65]. Adhesion and toughness of the NS diamond films produced from $N_2/CH_4/H_2$ plasmas with high CH_4 concentration were investigated by indentation testing up to 150 kg load with a 1/8 inch diameter tungsten carbide ball [26]. This results in significant

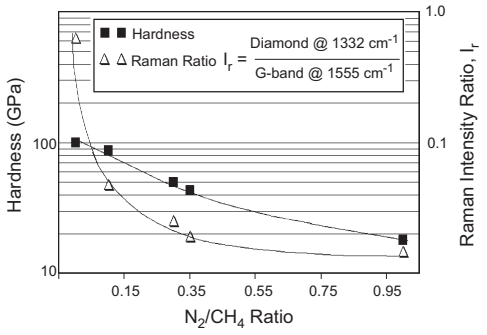


Figure 7. Nanoindentation hardness and Raman intensity ratio ($I_r = I_{1332}/I_{1555}$) vs N_2/CH_4 ratio. The hardness value for the N_2/CH_4 ratio of 0.0 was not measured but was estimated as 100 GPa (a value representative of high phase purity crystalline diamond). Reprinted with permission from [18], S. Catledge et al., *J. Appl. Phys.* 91, 5347 (2002). © 2002, American Institute of Physics.

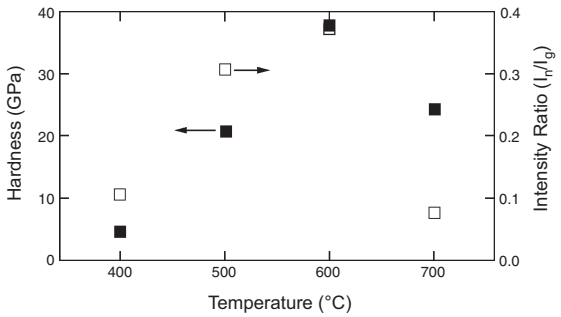


Figure 8. Plots of hardness and Raman intensity ratio of nanocrystalline diamond (I_n) to graphitic G band (I_g) as a function of substrate temperature for films deposited using the biased enhanced growth technique. Reprinted with permission from [58], T. Sharda et al., *Appl. Phys. Lett.* 77, 4303 (2000). © 2002, American Institute of Physics.

plastic deformation of the film as well as the underlying substrate and can provide a semiquantitative measure of film adhesion and toughness [62, 66]. An optical micrograph of the indentation craters is shown in Figure 9 for indentation loads of 294, 588, 980, and 1470 N. No delamination of the film was observed even for testing at the highest load. However, scanning electron microscopy revealed the presence of circumferential microcracks in the film around the edge of the indents. The strain to cause microcracking was estimated to be as high as 1.9%, which represents a significant improvement in toughness over other ceramic coatings. The cracks clearly did not result in catastrophic failure of the coating but are stable and likely responsible for the observed ductility. Similar tests performed on conventional diamond films (with micrometer-size grains) resulted in complete film spallation at the lowest loads used [25]. These results suggest that NS diamond films can be tailored with excellent adhesion on metallic surfaces and have hardness ranging from 10 (medium-hard) to 100 GPa (superhard).

In addition to high hardness and good adhesion, low coefficient of friction is necessary for many coating applications in which surfaces are in sliding contact with each other. Extensive work by Erdemir et al. [48, 67] to characterize friction and wear behavior of nanocrystalline diamond films grown in Ar–C_{60} and Ar–CH_4 plasmas has shown very low friction coefficients (0.04 in dry N_2; 0.1–0.15 in air) and wear rates, essentially comparable to those of natural diamond. The friction fluctuated substantially when tested in dry N_2, especially during long duration tests. This was attributed to transformation of the diamond surface into sp^2-bonded amorphous carbon wear debris particles that get trapped at the sliding interface.

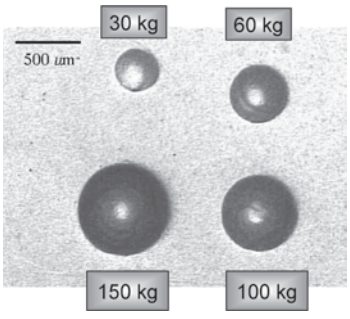


Figure 9. Optical micrograph showing a set of four indentations made on a nanostructured diamond-coated titanium alloy using a 1/8-in. diameter tungsten carbide ball indenter and loads of 30, 60, 100, and 150 kg. No delamination of the film was observed, indicating excellent film adhesion and toughness. Reprinted with permission from [148], S. A. Catledge et al., *J. Nanosci. Nanotechnol.* 2, 293 (2002). © 2002, American Scientific Publishers.

2.3.1. Wear of TMJ Implants

Wear of articulating surfaces is a major problem in orthopedic devices such as total hip and knee replacements. In dentistry, wear of the TMJ implant (shown in Fig. 1) is a concern. It has been estimated that more than 10 million people suffer from TMJ-related disorder symptoms in the United States alone [68]. The human TMJ, as diagrammed in Figure 10, has proven to be difficult to replace with a reliable artificial implant [69, 70]. This is due in part to the sliding/rotating articulation of this joint [71], which requires a component geometry that is more complex than those found in most other artificial joints. In addition, TMJ implants are small and reside in close proximity to the eye, the ear, various nerves, and the brain with very little bone available for fixation. As a result, TMJ implant service lifetimes have historically been short compared to other common joint replacements, and implants have even been the subject of product recall.

Very little basic science has been performed on such fundamental issues as the articulation mechanics of the

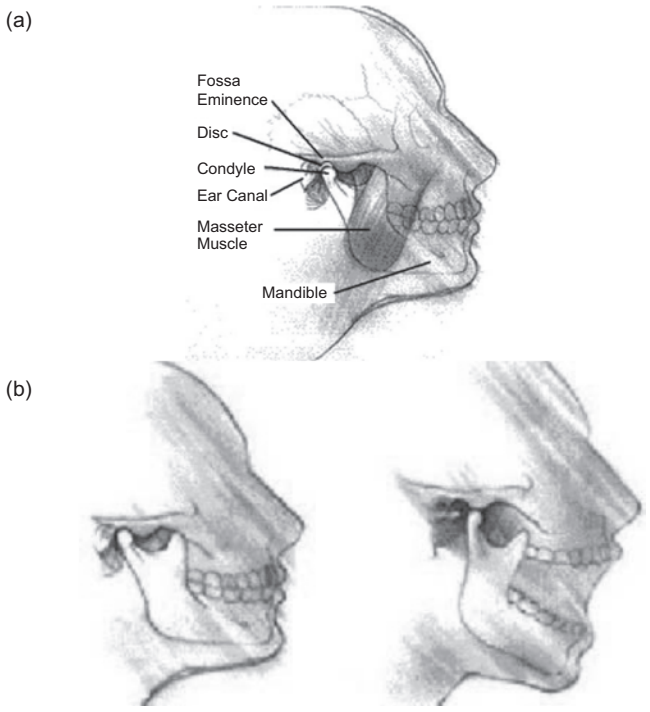


Figure 10. Diagrams showing TMJ anatomy and articulation motion. Notice the sliding/rotating articulation of the condyle and how it dislocates from the fossa eminence. Adapted with permission from [70], TMD: Temporomandibular Disorders, NIH Publication 94–3487, Bethesda, MD, 1994. © 1994, National Institutes of Health.

TMJ and disorder causes or progression. The basic structure of the TM joint is also poorly understood, as little research has been performed on typical TMJ loading forces, or on anatomical differences between genders and race. It is believed that the root cause of TMJ implant failure is a lack of understanding of the forces present during articulation [72]. Failure of joint prostheses typically occurs by one of three general conditions: (1) surgical error during placement, (2) postoperative infection, or (3) osteolysis [73] of the surrounding bone due to wear particle generation. Condition (3) results from a long-term failure mode, is least understood, and warrants intense investigation with regard to cause and prevention. During this failure mode, surface roughening of articulating implant components leads to increased wear particle generation from the polymer implant components that implant condyles articulate against [74, 75]. Polymer particles, in turn, cause osteolysis in the bone proximal to the implant [76, 77]. The implant then loosens leading to improper joint articulation, progressively severe implant wear, and accelerating bone loss. This third, long-term failure mode is dependent upon design of the implant, including the choice of materials used. Since long-term implant failure has its roots in the wear properties of the articulating components of the implant, an improvement of wear behavior should result in longer service lifetime for the implant and greater comfort and utility for the patient. It is hoped that a significant improvement can be achieved by coating implant articulation surfaces with a very hard, smooth layer of nanostructured diamond.

Nanostructured diamond films, 3 μm thick, were deposited on TMJ condyle simulants made from Ti-6Al-4V alloy. This alloy exhibits excellent mechanical properties [77] and superb resistance to electrochemical corrosion [78]. However, Ti-6Al-4V also shows poor wear resistance *in vivo* and the resulting titanium particles can cause bone resorption proximal to the implant [77]. For this reason, most metallic implant condyles are made almost exclusively of Co–Cr–Mo alloy today. Some patients, however, are especially sensitive to the nickel found in Co–Cr–Mo, and so a small number of implants are fashioned out of Ti-6Al-4V for these patients. NS diamond deposition on Ti-6Al-4V alloy has been studied extensively and is known to exhibit high hardness and low roughness [18, 26]. Since Ti-6Al-4V is acceptable as an articulating component in all terms but wear resistance, coating Ti-6Al-4V implant condyle components with NS diamond may result in a component that is acceptable for use as an implant in terms of wear resistance and biocompatibility.

Figure 11 shows images of two condyles, one with and one without the diamond coating, as well as a schematic of their dimensions. The coated TMJ simulants were evaluated for their wear resistance properties using a custom-built manibular movement simulator (MMS) developed by Dr. Camillo Machado at the University of El Bosque, Columbia. This is shown in Figure 12 [79]. This device is designed to accurately reproduce the normal articulation of a human jaw and is adapted to mount TMJ simulants in place of the natural TMJ anatomy. The MMS is a fiberglass reproduction of an actual human skull and so it closely mimics the mechanical movement and anatomical structure of a typical patient. The test is performed using a cyclic chewing

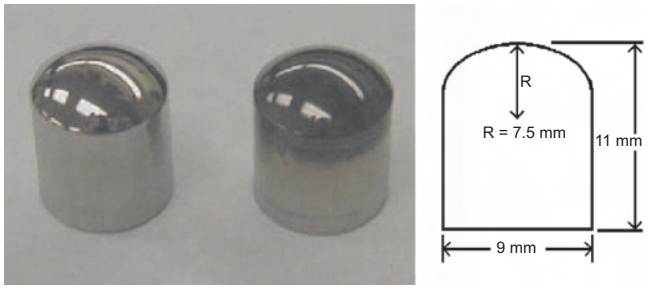


Figure 11. Ti-6Al-4V TMJ simulants uncoated (left) and coated with a 3 μm NS diamond (right), along with dimensions as they were measured from a commercially available implant.

motion of the mandible, with maximum chewing forces measured at the teeth of 200 N during testing. The NS diamond-coated TMJ stimulant was mounted in the anatomically correct temporomandibular position, opposite a 2.54-cm-diameter NS diamond-coated Ti-6Al-4V disk. The simulant was subjected to 500,000 loaded cycles in the MMS, which is equivalent to 4.4 years of typical use. It should be noted that the conditions of this test are considered to be much more extreme than those experienced by an actual implant because in this test the TMJ articulates against a flat plate in a point wear mode. An actual implant would articulate against a conformal fossa component that would increase the articulating surface area, thereby greatly decreasing the pressure at any point of the film in contact. Eventually, a NS diamond-coated fossa component will be fabricated to produce a more realistic TMJ couple and the wear test will be repeated.

A separate experiment was performed by Machado et al. [79] using uncoated Ti-6Al-4V condyles with NS diamond-coated Ti-6Al-4V flat disks. A control disk (uncoated Ti-6Al-4V) was also tested for comparison. The samples were subjected to 400,000 cycles at 1.2 Hz in the MMS. Surface profilometry of the wear tracks on the disks was performed to determine the extent of wear by measuring surface roughness. The rms surface roughness of the

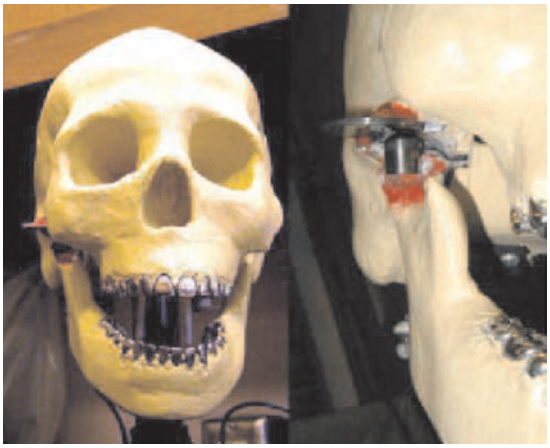


Figure 12. Mandibular movement simulator. The MMS (left image) is based on a composite replica of a human skull to replicate the natural motion of the TMJ. The image on the right shows detail of a TMJ condyle stimulant and wear test disk in position and prior to fixation.

uncoated control disk changed from 50.3 nm to 248.4 nm. In comparison the rms roughness of the NS diamond-coated disk changed from 61.5 to 122.7 nm. SEM also confirmed that damage to the uncoated Ti surfaces was much more extensive than for the diamond-coated surfaces.

2.4. *In Vivo* Biocompatibility

Lemons et al. have recently performed a study to assess the biocompatibility of NS diamond-coated Ti-6Al-4V alloy *in vivo* using animal models [80]. This work is complementary to the diamond-coated TMJ study in that it shows quite clearly that the animal models respond favorably to implantation of NS diamond materials on Ti-6Al-4V. These results show that NS diamond materials exhibit excellent biocompatibility and are therefore well suited as implant materials.

In the study, eight 7 mm diameter, 2 mm thick disks of Ti-6Al-4V alloy were coated with NS diamond to a thickness of 3 μm. The samples were seeded ultrasonically using 1–2 μm diamond powder and then diamond deposition was performed using a feedgas mixture of 500 sccm H_2, 88 sccm CH_4, and 8.8 sccm N_2. The uncoated side of each sample was then carefully polished to a mirror finish. The diamond coated samples were implanted into four New Zealand white rabbits, two samples in each rabbit. One sample was implanted into the tibial proximal condyle and one into the femoral distal condyle of the same leg on each rabbit. The rabbits were examined clinically and radiographically to ensure proper placement of the implant, and the rabbits were euthanized at times up to 8 weeks. The femoral and tibial condyles with implanted samples were excised and sectioned, five per sample, by the Exakt system. The Exakt system is a standard, commercially available process whereby the biological material is dehydrated with a series of alcohols and infused with xylenes to improve monomer penetration into the sample. The sample is then soaked in polymethylmethacrylate under vacuum at reduced temperature to allow the methylmethacrylate monomer to saturate the sample. The sample is then warmed to room temperature to allow the polymer to set, cut into thin sections with a diamond-blade band saw, and polished on a polishing wheel. The sections were then observed and photographed by an optical microscope.

Figure 13 shows a cross-section image of an implanted NS diamond-coated disk after retrieval. Living bone tissue had grown across both the NS diamond and Ti-6Al-4V surfaces. The retrieved samples showed no foreign body reactions, fibrous granulation tissue, or abnormal characteristics within the bone/marrow spaces. Both compact and trabecular bone are found to be directly associated with both the titanium alloy surface and the NS diamond-coated surface without the presence of fibrous tissue at the interface. Results were consistent throughout the retrieved samples. These findings show that NS diamond thin films perform at least as well as Ti-6Al-4V in terms of general and histological biocompatibility. Given the excellent biocompatibility properties of Ti-6Al-4V [81, 82], this finding identifies NS diamond thin films as a material well suited to long-term use in the human body.

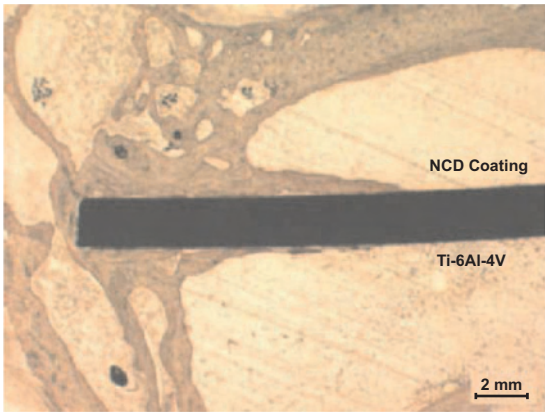


Figure 13. Cross section image of implanted NS diamond-coated disk. Living bone tissue has grown across both the NS diamond and Ti-6Al-4V surfaces. No immune response occurred in the marrow tissue adjoining the implant.

2.5. Corrosion Resistance

Biomedical implants and their components must operate in a corrosive environment, without maintenance or added lubrication, over service lifetimes that are expected to cover decades. As a result, there are stringent demands placed on these components in terms of high corrosion resistance and minimal wear particle generation. Nanostructured diamond films must meet these requirements in order to effectively perform as protective coatings. In terms of electrochemical resistance, they are uniquely suited to do so.

The use of diamond films as a protective coating for biomedical implants is an application with great promise. Currently, most articulating surfaces in human artificial joints are made of wrought Co–Cr–Mo alloy, which is chosen for its excellent mechanical wear and electrochemical corrosion properties. Much of the nonarticulating mass of implants is composed of Ti-6Al-4V alloy primarily due to this alloy's lower cost and excellent bone ingrowth behavior [83]. While Ti-6Al-4V would be preferable to Co–Cr–Mo in terms of bone/implant bulk modulus mismatch and cost, the poor mechanical wear properties of Ti-6Al-4V preclude its use in articulating components. The addition of a hard, smooth NS diamond coating to implant articulation surfaces may ameliorate this shortcoming. Previous work has shown that Ti-6Al-4V is especially well suited to the NS diamond deposition process and hard, smooth films over curved substrates made of this alloy can be reliably produced [84].

One aspect of NS diamond-coated Ti-6Al-4V that must be investigated is the electrochemical corrosion behavior of the coated alloy system. The electrochemical properties of diamond have been closely scrutinized, and in fact products based on diamond thin films have reached commercial production for the electrochemical research market. The primary attributes that draw attention to these films are their large potential window and excellent chemical stability [85–93], allowing for electrochemical measurements over a comparatively wide voltage range in caustic environments. Boron doping of diamond films can result in semimetallic electrical behavior, and Fausett et al. [93] report similar behavior for nanocrystalline diamond films grown in Ar-rich

gas mixtures without hydrogen. Studies have shown diamond electrodes to be stable in corrosive environments such as anodic polarization in acidic fluorine, and alkaline, neutral, and acidic chlorine [90]. To date, most electrochemical studies of diamond films have focused on defining their behavior for electronic applications and have not closely examined electrochemical corrosion behavior. However, Dowling et al. [94] used a scanning reference electrode technique to examine diamondlike carbon on AISI 316L stainless steel and found that coated samples could sustain a bias of 2400 mV without breaking down. By comparison, bare stainless steel control samples reached breakdown potential between 200 to 600 mV.

Biocompatibility tests performed by Mitura et al. [95] show that NS diamond coatings are very strongly resistant to chemical attack by Tyrode's solution (an aqueous solution with the addition of salts with properties similar to that in the human body) as well as to some HCL-based acids. Biotolerance was investigated by coating AISI 316L stainless steel disks with NS diamond and subcutaneously implanting them into the chest muscles of guinea pigs according to ASTM 98186 standards. It was found in the subcutaneous tissue, muscles, and bones that thin connective tissue capsules built from fibrocytes and collagen fibers were formed. In the wall of the capsule, no phagocyte reactions were observed and no products of corrosion were found. In addition, no pathomorphological changes were observed in the internal organs (liver, kidneys, and spleen). The study illustrated that NS diamond coatings exhibit good biotolerance and biocompatibility and effectively protect against corrosion and metalotoxemie.

Ti-6Al-4V is considered to be an excellent material for biomedical applications in terms of corrosion resistance due to its protective native oxide [96–98]. From a corrosion standpoint any surface treatment would be considered well suited for biomedical use if it were to preserve the corrosion properties of the alloy, much less improve upon it. To test this supposition, experiments were performed on passivated and NS diamond-coated Ti-6Al-4V substrates [99]. The substrates were polished to an rms roughness of 20–30 nm and then passivated following ASTM specifications (F 86) for implant materials. Three samples were coated with the NS diamond films to a thickness of 3 μm. The control and coated samples were subjected to potentiodynamic polarization in deaerated Hank's balanced salt solution. The counterelectrodes were platinum metal cylinders. All potentials were expressed with reference to a standard calomel electrode (SCE). The solution was purged with nitrogen gas from half an hour prior to immersion of the sample to the end of the test. The test procedure consisted of monitoring the rest potential for 1 hour, followed by generating cathodic and anodic polarization curves. Data were generated by conducting a forward scan from 100 mV more cathodic than the corrosion potential (E_{corr}) to a threshold anodic current density, and then reversing until the open circuit potential was reached. Tafel extrapolation and Stern–Geary currents were used to calculate the corrosion current density (I_{corr}) in A/cm^2 at the E_{corr}. The E_{corr} and the I_{corr} values obtained for the two sample groups are presented in Table 1. The E_{corr} values for the control samples were significantly more active (negative) than those exhibited by the NS diamond-coated samples. The noble E_{corr} values exhibited by the coated samples must have resulted from the protection provided by the insulating NS diamond film. I_{corr} values for the NS diamond-coated samples were an order of magnitude lower than those exhibited by the Ti-6Al-4V control samples. These E_{corr} values were also significantly more noble and were two orders of magnitude lower than those exhibited by a passivated Co–Cr–Mo alloy implant bearing surface evaluated in an earlier study under similar experimental conditions [100]. It should also be noted that the NS diamond-coated samples exhibited a significantly larger range of passive behavior than the control Ti-6Al-4V samples.

Table 1. Potentiodynamic polarization test results for NS diamond-coated Ti-6Al-4V in deaerated Hanks salt solution.

Sample	E_{corr} (mV vs SCE)	I_{corr} (nA/cm^2)
Ti-6Al-4V control	-96 ± 19	9.85 ± 5.86
NS diamond-coated	154 ± 13	0.93 ± 0.86

Source: Reprinted with permission from [148], S. A. Catledge et al., *J. Nanosci. Nanotechnol.* 2, 293 (2002). © 2002, American Scientific Publishers.

Figure 14 shows potentiodynamic polarization curves for passivated and NS diamond-coated Ti-6Al-4V, and Table 2 shows the calculated values of corrosion rate, I_{corr}, and open circuit potential (E_{OCP}) for the coated and control groups. No sign of electrochemical breakdown behavior all the way up to a potential of 1600 mV is observed for the uncoated Ti-6Al-4V material. This suggests that a very stable, well-adhered oxide coating is protecting the bulk Ti-6Al-4V from chemical attack. The NS diamond-coated samples also show no sign of breakdown behavior over the range of the experiment and the material becomes more noble as seen from the more positive open circuit potential (E_{OCP}). The corrosion rate of the system is correspondingly decreased approximately sevenfold. A one-way ANOVA test performed at $P = 0.05$ indicates that the mean corrosion rate values for NS diamond-coated Ti-6Al-4V and wrought Co–Cr–Mo are statistically indistinguishable. The mean corrosion rate of wrought Co–Cr–Mo was taken from a separate experiment

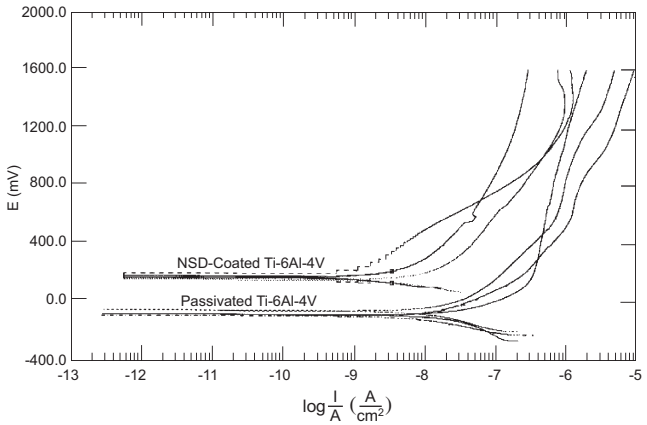


Figure 14. Potentiodynamic polarization curves for passivated and NS diamond-coated Ti-6Al-4V. Note the higher, more noble open circuit potential of the NS diamond-coated samples and the lack of breakdown behavior in any sample to 1600 mV.

Table 2. Measured values from potentiodynamic polarization for NS diamond-coated Ti-6Al-4V.

	Corrosion rate (mm/yr)	I_{CORR} (nA/cm^2)	E_{OCP} (mV)
Passivated Ti-6Al-4V	449.2E-6 +/− 269.9E-6	27.06 +/− 16.26	−93.4 +/− 19.2
NCD-coated Ti-6Al-4V	63.2E-6 +/− 87.8E-6	3.81 +/− 5.29	153.6 +/− 10.1
Wrought Co–Cr–Mo	24.8E-6 +/− 13.1E-6	94 +/− 0.79	−171 +/− 4.7

performed with the same apparatus and under the same conditions as this experiment [101].

The passivated, uncoated Ti-6Al-4V samples show a single curve in the Nyquist plot (Fig. 15) indicative of a protective oxide layer, as is typical for Ti-6Al-4V. The results are summarized in Table 3. The polarization resistance (R_P) value of 5.00×10^6 Ω measured from these data is very high and in agreement with the good corrosion resistance attributed to this alloy. In the case of the coated samples, the Nyquist plot shows a pair of impedance curves. These data are consistent with the behavior of microcrystalline, nanocrystalline, and single-crystal diamond thin films and are regarded as an intrinsic property of diamond [102, 103]. The Nyquist curves are also in agreement with the equivalent circuit proposed by van de Lagemaat et al. [103]. We believe that the native oxide does not survive the MPCVD deposition process as a result of exposure to hydrogen plasma during the thin film deposition step and so the oxide does not play a role in electrochemistry of the NS diamond-coated samples. Overall, the potentiodynamic polarization and electrochemical impedance spectroscopy data indicate that the film is an electrochemically protective material with high polarization resistance.

3. NANOSTRUCTURED HYDROXYAPATITE COATINGS

Some of the most chemically stable biomaterials used are ceramics. The surface and bulk chemistry can be controlled to tailor strength, hardness, wear resistance, and thermal expansion to each application individually. Along with the natural ability to promote tissue in-growth via a porous surface, the high compressive strength of ceramic materials lends itself for use as implants. Because of these properties, ceramics have been used as crowns, bridges, and false teeth in dental applications, in mandibular jaw reconstructions, as orthopedic load-bearing hip prostheses, and as coatings for porous tissue in-growth.

Hydroxyapatite $[Ca_{10}(PO4)_6(OH)_2]$ is the primary ceramic coating material used for both dental and orthopedic implant devices. HA is very similar to the mineral or inorganic portion of bone and has been shown to have good osteoconductive properties (i.e., to be conducive to bone formation around the implant). However, the low toughness (0.8–1.2 MPa $m^{1/2}$) and low flexural strength (<140 MPa) of bulk synthetic HA material precludes its use as a load-bearing implant [104]. Therefore, its primary use is either in the form of small filler particles for placement in bone-deficient areas or as coatings on stronger implant materials such as titanium and cobalt–chromium alloys. Calcium phosphate coatings (usually containing about 70% HA) are commonly deposited by the plasma spray process onto endosseous (i.e., into bone) dental implants in order to speed up formation of bone around the devices, thus allowing earlier stabilization in a patient [105]. Studies have shown that in cases of poor bone quality in a dental patient's implant site, a coating of HA is especially desirable, and coated devices are usually preferred over uncoated titanium implants. HA coatings have gained wider acceptance in dentistry as compared to orthopedics, but the potential for increased use in both fields is high if superior coatings can be developed.

The importance of the nanostructure of HA coatings for biomedical devices has not been realized until the last 10 years or so, as thin film deposition processes have been adapted from the electronic industry and employed in an attempt to produce superior HA implant coatings. Nanostructure processing of HA-based composites has allowed for the design of structural and surface features inspired by the architecture of bone. It has allowed chemical homogeneity and microstructural uniformity to be achieved for HA so that fully dense bioceramics can be generated at low sintering temperatures with a significant reduction in flaw size. HA powders can be easily densified to greater than 98% theoretical density, resulting in an ultrafine-grained (125 nm) compact that exhibits superior compressive and

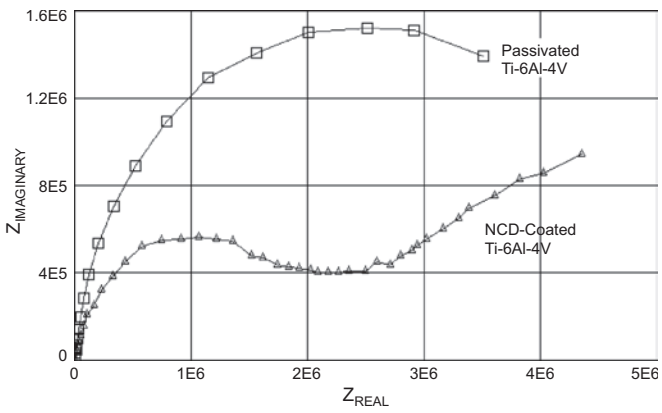


Figure 15. Nyquist plot of EIS data mean values for passivated and NS diamond-coated samples. Polarization resistance (R_P) and maximum frequency (ω_{MAX}) values taken from these data are reported in Table 3.

Table 3. Polarization resistance (R_P) and maximum frequency (ω_{MAX}) values for NS diamond-coated Ti-6Al-4V as measured from EIS data.

	R_P (Ω)	ω_{MAX}
Passivated Ti-6Al-4V	5.00E + 06	9.42E + 06
NS diamond-coated Ti-6Al-4V		
First curve	2.07E + 06	3.14E + 06
Second curve	1.24E + 07	7.85E + 06

bending strengths and fracture toughness. The nanometer-sized grains and high volume fraction of grain boundaries in nanostructured HA can increase osteoblast adhesion, proliferation, and mineralization. Zirconia nanocrystals have been added to further toughen the HA matrix. The resulting nanocomposite can achieve greater than 98% theoretical density with grain size less than 125 nm by pressure- and temperature-assisted sintering. The incorporation of highly dispersed zirconia significantly enhances the fracture toughness and bending strength of the HA-based bioceramics for load-bearing implant applications.

Almost all commercial bioceramic coatings placed on dental implants at this time are plasma sprayed, but these coatings have numerous limitations, such as inherent fractures, porosity, and unpredictable chemistry. These relatively thick (40–80 μm) coatings are subject to fracture or fragmentation during insertion and service and have unpredictable rates of dissolution. One of the primary reasons for the lack of control with the use of the plasma spray process is that extremely high temperatures must be employed in melting the starting powder (in excess of 10,000 °C), thus making it difficult to obtain the proper chemistry and structure of the resulting coating [106, 107].

Although calcium phosphate exhibits desirable properties such as biocompatibility, osteoconductivity, bioactivity, and direct bonding to bone interface, the bulk ceramic lacks the fracture toughness and flexural strength necessary for most load-bearing applications. However, the potential for HA coatings on metallic implants has resulted in a recent surge in mechanical property studies for this material [98–113]. Figure 16 shows an example of a commercially pure titanium implant screw (10 mm long, 3.75 mm diam) coated with 50 μm of plasma-sprayed HA. The HA coating extends from below the head of the screw to the insert-end base and is observed as an off-white color. X-ray diffraction confirmed the crystalline component of the coating to be hydroxyapatite [$Ca_{10}(PO_4)_6(OH)_2$] with hexagonal structure, space group $P6_3/m$, and measured lattice parameters $a = 9.393 \pm 0.003$ Å and $c = 6.885 \pm 0.003$ Å [114]. Nanoindentation testing was performed on the screw to yield measurements of coating hardness and Young's modulus. Hardness and Young's modulus measurements can provide a useful comparison to the mechanical properties of the surrounding bone and underlying implant. Figure 17 shows hardness and

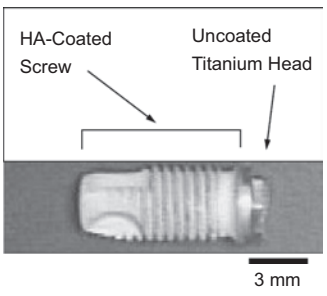


Figure 16. Photograph of a 10-mm-long titanium implant screw with 50-μm-thick hydroxyapatite coating made by the plasma spray process. The coating does not cover the head of the screw. Reprinted with permission from [148], S. A. Catledge et al., *J. Nanosci. Nanotechnol.* 2, 293 (2002). © 2002, American Scientific Publishers.

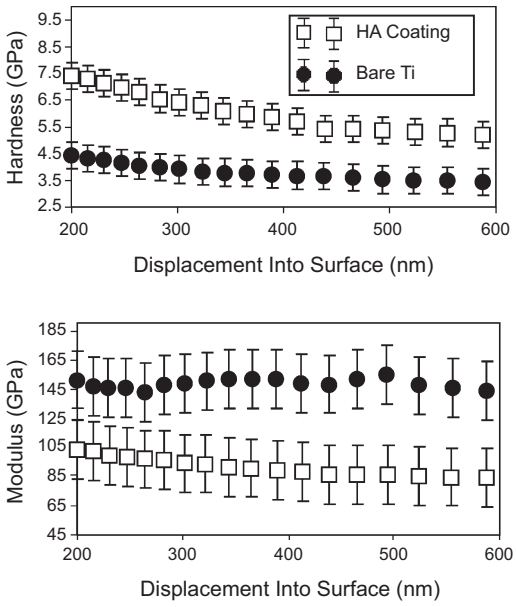


Figure 17. Nanoindentation hardness and Young's modulus vs indentation depth for the plasma sprayed hydroxyapatite coating deposited on the titanium implant screw shown in Figure 17. For comparison, similar measurements were also obtained from the uncoated titanium surface. Reprinted with permission from [148], S. A. Catledge et al., *J. Nanosci. Nanotechnol.* 2, 293 (2002). © 2002, American Scientific Publishers.

Young's modulus for the HA screw coating as well as that obtained from the uncoated titanium surface; each data set is plotted with error bars. This error represents the standard deviation from 12 indents at that depth. The average values of hardness and modulus (6.1 ± 0.4 and 89 ± 6 GPa, respectively, at 400 nm depth) represent critical mechanical property measurements for HA coatings that must withstand large abrasion forces during implant insertion.

Concerns about the lack of control of microstructure and the relatively low bond strength with the metallic implant substrate cause some scientists and clinicians alike to reject the use of plasma sprayed HA-coated implants, in spite of the apparent benefits of these coatings. The average grain size of HA crystals in these coatings is 15–25 μm, and each grain is surrounded by amorphous regions resulting from the rapid cooling of the semimolten ceramic material [115]. The generation of unwanted calcium oxide and other more dissolvable calcium phosphate crystalline phases (such as tricalcium phosphate) is also quite common with the plasma spray process. The size of the grains and the percent of amorphous material and non-HA crystalline phases influence the dissolution rate. Studies have indicated that a 46% crystalline HA plasma sprayed coating dissolves three times as rapidly as a 75% crystalline HA coating [116, 117].

The size of the grains for plasma sprayed coatings is based on the initial size of the starting powder, which is typically in the range of 20 to 40 μm (average about 35 μm). It is not possible to use smaller particles as feedstock in the plasma spray process because they will volatilize and result in significant compositional changes. The grains or crystals in the deposited coating are the remaining cores of the starting powder that did not dissolve in the argon gas plasma. The

rapid cooling of the semimolten powder leaves little opportunity for microstructural control during the plasma spraying process, and postdeposition heat treatment has been found to be of little value in optimizing the microstructure of the coatings. Because of this lack of control of the coatings on either the micro- or nanostructural level, a coating with optimal dissolution properties has not been achieved using plasma spraying or thermal spray processes in general. Incremental advances in HA coatings require that a different coating method be used that will allow control of the crystalline structure on a nanostructural level.

Previous studies at the University of Alabama at Birmingham (UAB) have demonstrated that several alternative coating techniques have the potential for producing calcium phosphate coatings that are superior to the existing commercial plasma sprayed coatings. The two methods shown to be most successful in producing HA coatings with controlled chemistry and structure are ion beam sputtering and pulsed laser deposition, both of which are described in previous publications [118, 119]. Ion beam sputter deposited coatings were shown to possess superior bond strength and density as compared to plasma sprayed coatings [120, 121]. Figure 18 shows the crystalline structure of a sputter deposited calcium phosphate coating. Sputtering can produce nanocrystalline coatings when followed by an appropriate heat treatment [122]. The primary value of both of these vacuum chamber processes is that nanosize grains can be formed because the coatings are essentially built up in small increments, or on an atom-by-atom basis if desired. In both of these coating techniques, material is removed from a pressed/sintered HA target by bombardment with a beam. This is usually an argon ion beam in the case of ion beam sputtering, or by a laser in the pulsed laser deposition (PLD) process. In both techniques the coatings can be built up slowly, thus permitting the precise tailoring of deposition conditions to grow nanosize crystals rather than the microsize grains inherent in the plasma spray process.

Desirable aspects of thin, HA coatings are high density, high abrasion resistance (important during implant insertion), the ability to release a small amount of calcium and phosphate ions to stimulate bone growth, and sufficient durability to remain on the metallic implant until bone apposition has taken place [123]. Of course, the ability to enhance bone formation around the implant is the primary reason for using any type of calcium phosphate coating, and thin film coatings must accomplish this objective in the human body prior to dissolution. Initial results using nanocrystalline bioceramic coatings are encouraging, as thin calcium phosphate coatings deposited by radio frequency sputtering were shown to possess sufficient durability and to provide stronger implant attachment to bone as compared to uncoated titanium implants [124].

When comparing the various coating methods, laboratory studies at UAB have indicated that the PLD technique is overall the most promising means of placing nanocrystalline HA coatings on titanium and other metallic implant materials. For calcium phosphate coatings deposited by ion beam sputtering, control of the composition and grain size was found to be slightly more difficult as compared to PLD, and microstructural analyses showed that the sputtered coatings were less dense and less crystalline on deposition than those obtained by PLD [125].

Typically, either of the vacuum deposition processes, sputter coating or PLD, produces amorphous calcium phosphate films (when using HA targets) that do not possess sufficient resistance to dissolution. Crystallization is necessary for these coatings to stay on the implant long enough to provide a useful benefit, and that can be accomplished by either closely controlling conditions in the vacuum chamber during deposition or by the use of a postdeposition heat treatment. Typical processing conditions for PLD include the use of a starting vacuum in the range of 2×10^{-5} Torr, a laser energy of about 5 J/cm^2, the introduction of a water vapor/argon mixture during deposition (to replace lost OH ions from the structure), and the use of a substrate heater to provide the thermal energy necessary to grow crystals [126]. Microstructural analyses showed that the PLD coatings produced at UAB using these parameters were very dense and of a much higher quality than those obtained by plasma sprayed HA coatings. The ability to bind specific proteins was demonstrated to be higher than uncoated specimens, thus indicating the potential for more rapid bone attachment to the implant [127]. Further research will be necessary to determine the precise role of nanostructure in influencing properties such as dissolution, abrasion resistance, bond strength, and *in vivo* performance of bioceramic HA coatings for dental and orthopedic implants.

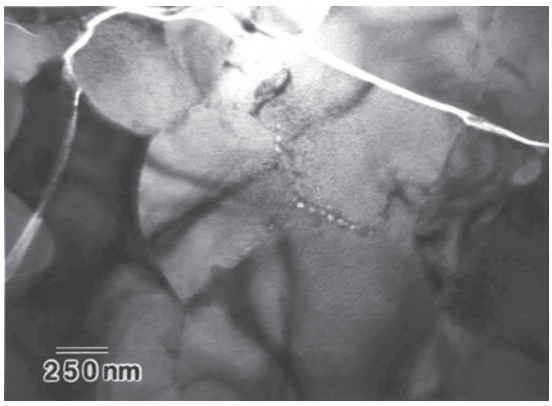


Figure 18. TEM image showing crystalline structure of a sputter deposited calcium phosphate coating, indicating that sputtering can produce nanocrystalline coatings when followed by an appropriate heat treatment. Reprinted with permission from [148], S. A. Catledge et al., *J. Nanosci. Nanotechnol.* 2, 293 (2002). © 2002, American Scientific Publishers.

4. NANOSTRUCTURED METALLOCERAMIC COATINGS

Total joint replacements have provided a viable option for restoring function to individuals whose joints have been afflicted by arthritis, joint fractures, bone tumors, and even trauma. Presently, total hip joint replacements consist of an ultrahigh molecular weight polyethylene (UHMWPE) acetabular cup in contact with a metallic femoral head, commonly fabricated from a wrought Co–Cr–Mo alloy. A metallic-plastic combination has been used for the past

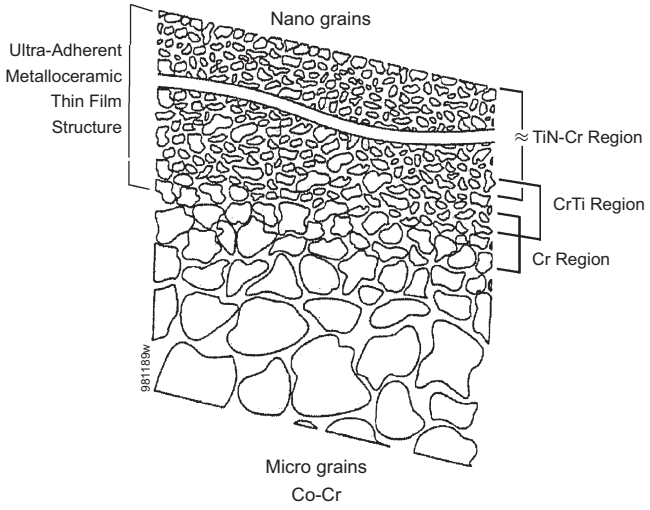


Figure 19. Schematic of functionally graded, nanocrystalline metalloceramic coating deposited by ion-beam assisted deposition. Adapted with permission from [101], S. Woodward, A Novel Metalloceramic Coating for Reduced UHMWPE Wear in Biomedical Devices, Master's Thesis, University of Alabama at Birmingham, 2002.

30 years; however, it is not without clinical shortcomings. Wear of the UHMWPE has been cited as a dominant factor limiting the long-term success of these devices [128–130]. To increase longevity of orthopedic implants, researchers have evaluated potential alternative materials for the articulating surfaces.

The wear process has been shown to be dependent upon the hardness and the roughness of the counterface. Slight changes in roughness can produce a large increase in UHMWPE wear rate [131–133]. A harder, more scratch-resistant surface will improve the resistance to third body wear [134]. Noting the importance of a hard surface layer for tribological purposes, various surface hardening processes have been evaluated for orthopedic devices. These include nitrogen ion implantation [135, 136], plasma ion nitriding, and physical vapor deposition (PVD) of various coatings [137]. All have met with limited success. One of the most detrimental problems in coating applications is inadequate adhesion to the substrate. Adhesion is defined as the joining of superficial layers of two different bodies brought into contact through the presence of attractive forces [138]. There are three primary chemical bonds that materials possess: metallic, ionic, and covalent. Materials with covalent bonding are hard, are strong, and have high wear resistance. Metallic materials are tough and ductile, while ionic materials are hard and stable. Ceramic materials typically have a combination of covalent and ionic bonding. This difference in bonding leads to adhesion problems when depositing a ceramic coating on a metallic substrate. Intermediate layers are often used to increase adhesion. For example, for the deposition of diamondlike carbon (DLC) on titanium, a Ti–TiN–TiC–DLC multilayer structure was used to improve adhesion [139]. Hubler et al. [140] used a titanium intermediate layer to improve adhesion of TiN coating on a silicon substrate. The idea is to gradually move from metallic bonding at the substrate to primarily covalent bonding at the surface.

A novel, functionally graded nanocrystalline metalloceramic coating for Co–Cr–Mo components has been developed by Spire Corporation (Bedford, MA), which may potentially reduce UHMWPE wear in orthopedic devices without the common adhesion problems evident in coating applications. This multilayer coating deposited by ion beam assisted deposition (IBAD) has metallic initial layers and a hard, nanocrystalline ceramic outer coating (Cr/CrTi/CrTiN). A schematic of the graded coating structure is shown in Figure 19. A (Cr/CrTi) metallic layer at the interface will increase adhesion to the cobalt chrome substrate, while the surface layer (CrTiN) is primarily covalent in nature and will enhance scratch and wear resistance of the coating. The chromium metal layer increases adhesion between layers and increases strength of the coating by serving as a barrier to crack formation and propagation. These coatings have three primary advantages over conventionally deposited ceramic films: nanocrystalline structure with grain size about 10 nm, functionally graded layer whose bonding gradually changes from metallic to covalent with increasing ceramic material toward the surface, and inclusion of nanocrystalline metal to enhance strength and toughness (as a barrier to crack formation and propagation). The coatings will have metallic adherence to metals, but with significantly better wear resistance and a lower friction coefficient. The benefits of the film in orthopedic applications will be substantially enhanced surface hardness and wear resistance, enhanced adhesion compared to ceramic coatings, conformal surface which maintains a smooth finish, and retention of surface finish *in vivo*, leading to lower UHMWPE wear. An example of the smooth surface exhibited by the metalloceramic coatings is illustrated in the atomic force microscopy (AFM) images in Figure 20 for both the coated and uncoated Co–Cr–Mo disks. The AFM images are very similar for both samples. The coated sample had an rms surface roughness value of 11.1 nm, while the uncoated sample had an rms value of 11.5 nm.

Figure 21 shows the thin film XRD pattern of the metalloceramic CrTiN coating with peaks indexed to a cubic sodium chloride structure. The coating shows a preferred (200) orientation with [200] direction being aligned perpendicular to the substrate surface. The diffraction peaks are broad as expected for a nanocrystalline material and the grain size was calculated from the Scherrer equation to be 12 nm.

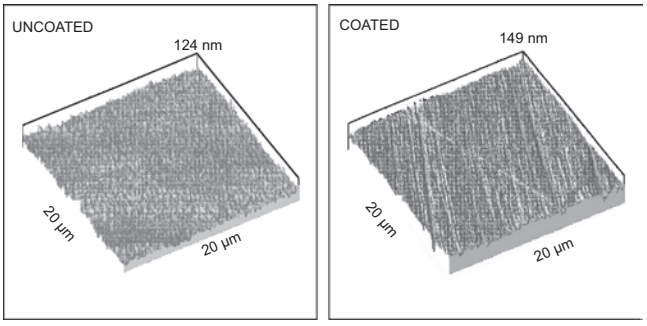


Figure 20. Atomic force microscope images of metalloceramic-coated and uncoated Co–Cr–Mo samples. Reprinted with permission from [148], S. A. Catledge et al., *J. Nanosci. Nanotechnol.* 2, 293 (2002). © 2002, American Scientific Publishers.

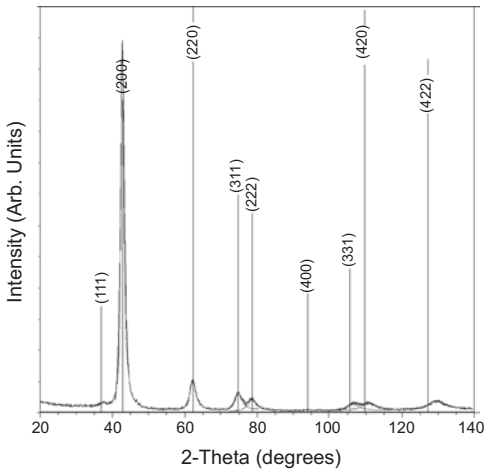


Figure 21. Thin film XRD pattern of the CrTiN metalloceramic coating. The miller indices (*hkl*) are indicated for the cubic sodium chloride structure. The fitted value for the lattice parameter is 4.2169 ± 0.0035 Å.

The bars indicate a least squares fit to the cubic structure resulting in a best fit lattice parameter of 4.2169 ± 0.0035 Å. Table 4 shows *d*-spacings of the fitted peaks with the percent error between observed and calculated values. It can be seen that the fit does not provide an excellent match to the weak (111) reflection or to the higher index peaks (331), (420), and (422). Table 5 shows the lattice parameter for bulk TiN as well as for experimentally determined TiN films prepared by different methods including PVD, N^+-IBAD, annealed N^+-IBAD, and Xe^+-IBAD. These values are all higher than our experimental lattice parameter for the metalloceramic film, which lies between that of CrN and TiN having lattice parameters of 4.140 and 4.244 Å, respectively. Therefore, since the XRD pattern provides a closer fit to TiN but with a lower lattice parameter than bulk TiN, we suggest that the surface layer is a tertiary Ti–N–Cr disordered solid solution which is predominantly TiN, but with some Cr atoms substituted for Ti.

Figure 22 shows nanoindentation hardness and Young's modulus vs indentation depth for the metalloceramic coating as well as for the uncoated Co–Cr–Mo substrate. The Young's modulus of the coating shows a drop from 325 GPa at 150 nm depth to 285 GPa at a depth of 600 nm. This drop

Table 4. Comparison of observed and calculated *d*-spacings for XRD peaks indexed for metalloceramic coating.

Miller index	Observed *d*-spacing (Å)	Calculated *d*-spacing (Å)	% Error
111 (very weak)	2.3957	2.4503	2.223
200	2.1128	2.1220	0.434
220	1.4930	1.5005	0.500
311	1.2695	1.2796	0.789
222	1.2171	1.2251	0.653
331	0.9618	0.9736	1.212
420	0.9361	0.9490	1.359
422	0.8518	0.8663	1.674

Note: The calculated *d*-spacings are based on a lattice parameter $a = 4.2169$ Å.

Table 5. Comparison of lattice parameters for TiN coatings prepared from different methods. The value for bulk TiN is also shown.

Preparation Method	PVD[a]	N^+-IBAD[a]	N^+-IBAD[a] treated at 850 °C	Xe^+-IBAD[a]	Bulk Tin
Lattice parameter (Å)	4.221	4.227	4.235	4.240	4.244

[a] From [147].

is likely due to interaction of the long-range elastic strain field with the underlying substrate having a lower modulus of 240 GPa. The hardness measurement is not as sensitive to this substrate effect since the region of plastic deformation is of shorter range. The measured average hardness of the metalloceramic coating is 27 GPa while that of the uncoated Co–Cr–Mo substrate is 10 GPa. The coating hardness is similar to values reported for TiN coatings [141]. For comparison, nanocomposite hard coatings such as those based on nanocrystalline TiC grains embedded in an amorphous carbon matrix have been engineered to have 32 GPa hardness and an extremely high plasticity of 40% (in nanoindentation deformation) [142]. As shown in the load vs displacement curve in Figure 23, the metalloceramic coating shows a remarkably high plasticity of 55%. Such plastic compliance is not typical of a hard ceramic coating and represents more than a fivefold increase over that observed for 60–70 GPa hard DLC coatings [143, 144]. The higher plasticity of the metalloceramic coating is attributed to the graded metallic-to-covalent bonding structure. The chromium subsurface metallic bonding promotes toughness within the coating while the outer primarily covalently bonded TiN–Cr layer enhances the surface hardness. The combination of a graded metallic–ceramic nanocrystalline film with metallic interfacial adhesion and ceramic surface hardness makes them unique and very attractive for implant applications.

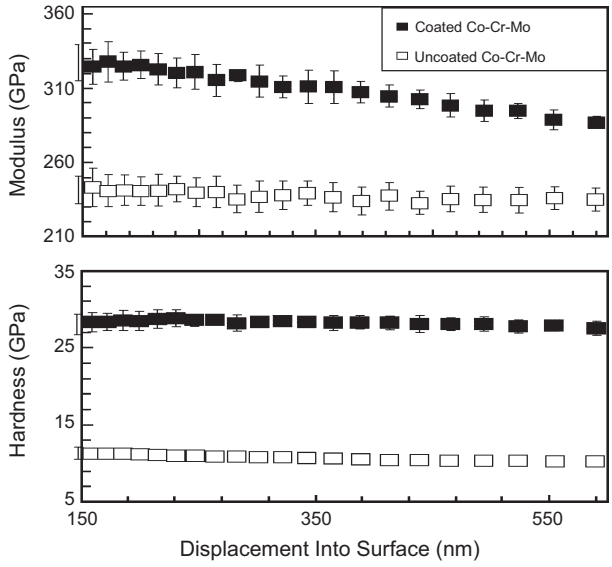


Figure 22. Nanoindentation measurements of Young's modulus and hardness for the metalloceramic coating and uncoated Co–Cr–Mo substrate.

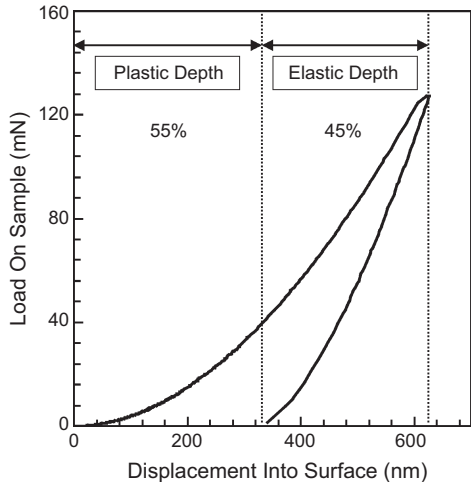


Figure 23. Nanoindentation load vs displacement curve for metalloceramic coating showing the large plasticity (55%).

The functionally graded and nanocrystalline nature of the metalloceramic coating is also expected to improve adhesion at the film/substrate interface. The bonding at this interface is completely metallic since a chromium layer was deposited as the first processing step. The chromium should serve as a barrier to crack formation and propagation. In addition, the nanosize grains should enhance the strength and toughness near the interface as compared to an interface having conventional grain size and/or having a more abrupt transition in bonding. In order to evaluate adhesion of the metalloceramic coating, scratch testing was performed. The scratch tip was a hemispherical diamond with 125 μm radius. Figure 24 shows the normal force and acoustic signal vs scratch distance produced during a 10 mm scratch while ramping the normal force up to 60 N. A critical force of 39 N was established, upon which abrupt changes in the acoustic signal indicated cracking/decohesion of the coating. This was confirmed by detailed microscopic investigation of the scratch. Several scratches were performed to obtain an average critical force of 44 ± 5 N. The inset in the figure shows an optical micrograph of the scratch segment corresponding to the time near the critical event. The coating shows only very localized spalling near the edge of the scratch. No catastrophic cracking or delamination occurred, indicating excellent interfacial adhesion.

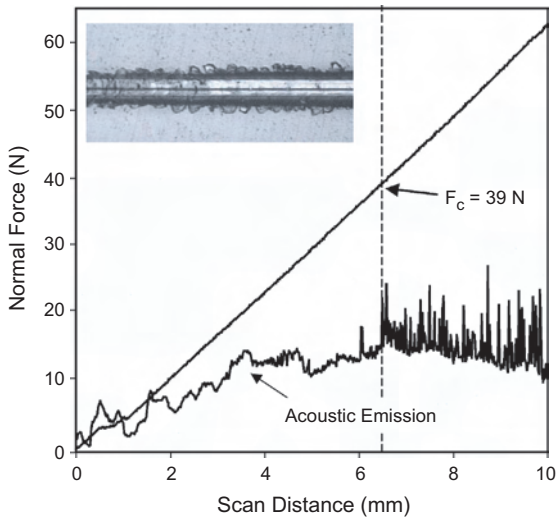


Figure 24. Plot from scratch adhesion test showing normal force and acoustic signal during a 10 mm scratch on metalloceramic coating. The inset shows an optical micrograph of the scratch segment corresponding to the time near the critical event.

The corrosion behavior of potential biomaterials must be addressed due to the corrosive environment of the human body. Being ceramic in nature one would expect a greater corrosion resistance with the metalloceramic coating (Cr/CrTi/CrTiN). This hypothesis has been affirmed through potentiodynamic polarization testing. Figure 25 shows polarization curves for metalloceramic-coated and uncoated control samples in deaerated neutral pH Hanks salt solution at 37 °C. The polarization curve for the coated sample is displaced a log decade to the left of the control sample, illustrating the order of magnitude lower corrosion rate of the coated sample. The breakdown potential (E_{bd}) of the coated sample (703 mV vs SCE) was significantly higher than the uncoated control sample (469 mV vs SCE). Also, on reversal of the polarization scan after breakdown, the coated sample exhibited a reduced current density (no hysterisis) while the uncoated sample exhibited a higher current density until open circuit potential was reached (large hysterisis). This indicates that pitting and crevice corrosion will propagate after breakdown on the uncoated substrate in comparison to the coated sample. Also, the coated sample may be able to repassivate any surface damage due to mechanical/electrochemical interactions. T-tests found that the corrosion rates and the breakdown potentials were significantly different between the coated and uncoated disks at $p < 0.05$.

Studies of the wear resistance of metalloceramic coatings are in progress. Previous investigators have demonstrated that a titanium nitride coating reduced UHMWPE wear. There was no delamination or penetration of the ceramic coating as found by Pappas et al. [145]. Likewise, in knee simulation studies conducted by Peterson et al. [146], titanium nitride coated Ti-6Al-4V femoral components showed

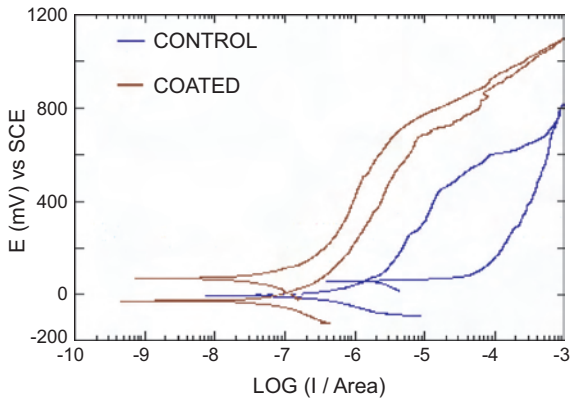


Figure 25. Overlaid representative polarization curves for control and metalloceramic-coated samples in deaerated neutral pH Hanks salt solution at 37 °C. Reprinted with permission from [148], S. A. Catledge et al., *J. Nanosci. Nanotechnol.* 2, 293 (2002). © 2002, American Scientific Publishers.

no scratches and retained a polished surface similar to that of cobalt–chromium–molybdenum components. There was no coating breakdown throughout testing. The metalloceramic coating reviewed here has exhibited promising characteristics. Ultimately, wear tests will predict if this coating should be further considered as a possible alternative in hip devices.

5. CONCLUSIONS

Although the wear and corrosion resistance of the ceramic coatings discussed in this chapter are promising for use in biomedical implant applications, further studies are needed to confirm their long-term safety, biocompatibility, and performance in use. Nevertheless, the benefits of nanostructured diamond, hydroxyapatite, and metalloceramic coatings are obvious.

Nanostructuring of diamond can occur via several processing routes and results in a dramatic improvement in fracture toughness and adhesion, reduced friction and wear rate, and higher corrosion resistance. The mechanical properties of diamond films can easily be tailored over a wide range by the appropriate manipulation of processing parameters. The nanostructured diamond coatings provide an ideal testing ground for the global hypothesis for wear studies that a particular "superhard" or "medium-hard" nanocrystalline diamond coated metal surface will perform better in laboratory wear tests and *in vivo* and may result in next generation of implant devices. The studies show that it is possible in the case of nanostructured diamond coatings to change the hardness and roughness parameter independently.

Nanostructuring of hydroxyapatite coatings by ion beam sputtering or pulsed laser deposition appears to allow for better control of chemistry and microstructure when compared to the more conventional plasma spray process. This, in turn, results in higher coating density and bond strength. The nanocrystalline coatings were shown to possess sufficient durability and to provide stronger implant attachment to bone as compared to uncoated titanium implants on metals. They promote better attachment to the bone while optimizing abrasion resistance, bond strength, and the dissolution rate. The nanometer-sized grains and high volume fraction of grain boundaries in nanostructured hydroxyapatite have been found to increase osteoblast adhesion, proliferation, and mineralization. Further research will be necessary to determine the precise role of nanostructure in influencing properties such as dissolution, abrasion resistance, bond strength, and *in vivo* performance of bioceramic hydroxyapatite coatings for dental and orthopedic implants.

Nanostructured metalloceramic coatings with a functionally graded structure from metallic (at the coating/substrate interface) to ceramic at the surface provide improved adhesion over conventional ceramic coatings on metals. The (Cr/CrTi) metallic layer at the interface will increase adhesion to the cobalt chrome substrate, while the surface layer (CrTiN) is primarily covalent in nature and will enhance scratch and wear resistance of the coating. The benefits of the film in orthopedic applications will be substantially enhanced surface hardness and wear resistance, enhanced adhesion compared to ceramic coatings, conformal surface which maintains a smooth finish, and retention of surface finish *in vivo*, leading to lower UHMWPE wear. These coatings show good corrosion resistance but further tests are needed to evaluate their wear-resistance potential for biomedical implants. Tables 6 and 7 provide a comparison summary of the properties and advantages/disadvantages for each NS coating described in this chapter.

Table 6. Summary of properties for each of the nanostructured diamond, hydroxyapatite, and metalloceramic coatings discussed in this chapter.

	NS diamond	NS hydroxyapatite	NS metalloceramic
Hardness (GPa)	85	6	27
Young's modulus (GPa)	700	89	305
Surface roughness (nm)	15	not measured	11

Source: Reprinted with permission from [148], S. A. Catledge et al., *J. Nanosci. Nanotechnol.* 2, 293 (2002). © 2002, American Scientific Publishers.

Table 7. Summary of the advantages and disadvantages for each of the nanostructured diamond, hydroxyapatite, and metalloceramic coatings discussed in this chapter.

	NS diamond	NS hydroxyapatite	NS metalloceramic
Advantages	• combination of extreme hardness, low surface roughness, and good fracture toughness	• promotes bone formation around implant • increases osteoblast adhesion, proliferation, and mineralization	• Combination of high hardness, very low surface roughness, and extremely good adhesion
	• good adhesion to titanium alloys • good corrosion resistance		• good corrosion resistance
			• conformal coatings
	• easily tailored mechanical properties		
Disadvantages	• poor adhesion to cobalt chrome and steel substrates • conformal coatings not well developed yet	• low toughness and flexural strength; not useful for load-bearing applications	• still in early state of development and testing; many properties remain unknown
		• control of chemistry and microstructure remains a challenge	

Source: Reprinted with permission from [148], S. A. Catledge et al., *J. Nanosci. Nanotechnol.* 2, 293 (2002). © 2002, American Scientific Publishers.

GLOSSARY

Chemical vapor deposition (CVD) Deposited species are formed as a result of chemical reaction between gaseous reactants at elevated temperature in the vicinity of the substrate. The solid product of the reaction is deposited on the surface of the substrate and is used to deposit films of semiconductors (crystalline and noncrystalline), insulators, as well as metals. Variations of CVD processes include atmospheric pressure CVD, low pressure CVD, and plasma enhanced CVD.

Hydroxyapatite A compound that is chemically similar to the mineral component of bones and hard tissues in mammals. It is one of few materials that are classed as bioactive, meaning that it will support bone ingrowth and osteointegration when used in orthopedic, dental and maxillofacial applications.

Metalloceramic coating A graded coating with a gradual transition from metallic bonding to ionic/covalent bonding typical of ceramics. The coatings will have metallic adherence to metals, but with significantly better wear resistance and a lower friction coefficient. The benefits of the film in orthopedic applications will be: substantially enhanced surface hardness and wear resistance, enhanced adhesion compared to ceramic coatings, conformal surface which maintains a smooth finish and retention of surface finish *in vivo*, leading to lower UHMWPE wear.

Orthopedics Orthopedics is the branch of medicine concerned with diseases, injuries, and conditions of the musculoskeletal system—relating to the body's muscles and skeleton, and including the joints, ligaments, tendons, and nerves.

Osteoblasts Cells which aid the growth and development of teeth and bones.

Tribology The branch of engineering that deals with the interaction of surfaces in relative motion (as in bearings or gears): their design and friction and wear and lubrication.

ACKNOWLEDGMENTS

We acknowledge support from the National Institute of Dental and Craniofacial Research, National Institutes of Health under grant 5R01D013952-02. We thank Spire Corporation (Bedford, MA) for the nanocrystalline metalloceramic samples for some of the studies described in this chapter. In addition, we thank Dr. Dieter Gruen of Argonne National Laboratory, USA and Dr. T. Soga of Nagoya Institute of Technology, Japan for allowing us to use results from their publications in this chapter. Finally, we thank Dr. Bill Lacefield, Dr. Jack Lemons, Dr. Krishna Venugopalan, Dr. Jones Papo, Dr. Camillo Machado, and Shanna Woodard for their useful discussions and contributions.

REFERENCES

1. A. M. El-Sherik, U. Erb, G. Palumbo, and K. T. Aust, *Scripta Metall. Mater.* 27, 1185 (1992).
2. A. A. Voevodin and J. S. Zabinski, *J. Mater. Sci.* 33, 319 (1998).
3. National Center for Health Statistics, 1994 Report.
4. S. Praemer and D. Furner, "Musculoskeletal Conditions in the United States," pp 125–141. AAOS, Rosemont IL, 1992.
5. P. Kumar, M. Oka, K. Ikeuchi, K. Shimizu, T. Yamamuro, H. Okumura, and Y. Kotoura, *J. Biomed. Mater. Res.* 25, 813 (1991).
6. D. N. Baird and W. J. Rea, *J. Nutritional Environmental Med.* 9, 209 (1999).
7. W. E. Shankland, in "TMJ; Its Many Faces," 2nd ed. Anadem, 1998.
8. W. von Bolton, *Z. Elektrochem*. 17, 971 (1911).
9. A. D. Kiffer, Tonawanda Laboratories, Linde Air Products Co., Synthesis of Diamond from Carbon Monoxide, 6 June 1956.
10. F. P. Bundy, H. T. Hall, H. M. Strong, and R. H. Wentorf, *Nature* 176, 51 (1955).
11. B. V. Deryagin and D. B. Fedoseev, *Sci. Am*. 233, 102 (1975).
12. B. V. Spitsyn, L. L. Bouilov, and B. V. Deryagin, *J. Cryst. Growth* 52, 219 (1981).
13. S. Matsumoto, Y. Sato, M. Kamo, and N. Setaka, *Jpn. J. Appl. Phys.* 21, 183 (1982).
14. S. A. Barnett, in "Physics of Thin Films" (M. H. Francombe and J. L. Vossen, Eds.), Vol. 17. Academic Press, Boston, 1993.
15. W. D. Sproul, *Surf. Coat. Tech.* 86/87, 170 (1996).
16. S. Veprek, *Thin Solid Films* 297, 145 (1997).
17. S. Veprek, P. Nesladek, A, Niederhofer, and F. Glatz, *Nanostruct. Mater.* 10, 679 (1998).
18. S. A. Catledge, J. Borham, M. Fries, Y. K. Vohra, W. R. Lacefield, J. E. Lemons, S. Woodard, and R. Venugopalan , *J. Appl. Phys.* 91, 5347 (2002).
19. N. Savvides and T. J. Bell, *J. Appl. Phys*. 72, 2791 (1992).
20. C. J. McHargue, in "Applications of Diamond Films and Related Materials" (Y. Tzeng, M. Yoshikawa, M. Murakawa, and A. Feldman, Eds.). Elsevier Science, Amsterdam, 1991.
21. J. J. Lander and J. Morrison, *Surf. Sci.* 4, 241 (1966).
22. J. C. Angus, N. C. Gardner, D. J. Poferl, S. P Chauhan, T. J. Dyble, and P. Sung, *Sin. Almazy* 3, 38 (1971); presented at International Conference on Applications of Synthetic Diamonds in Industry, Kiev, 1971.
23. D. V. Fedoseev, K. S. Uspenskaya, V. P. Varnin, and S. P Vnukov, *Izv. Akad. Nauk SSSR Ser. Khim.* 6, 152 (1978).
24. J. C. Angus and C. C. Hayman, *Science* 241, 643 (1991).
25. S. A. Catledge and Y. K. Vohra, *J. Appl. Phys.* 86, 698 (1999).
26. N. Toprani, S. A. Catledge, Y. K. Vohra, and R. Thompson, *J. Mater. Res.* 15, 1052 (2000).
27. Y. Namba, E. Heidarpor, and M. Nakayama, *J. Appl. Phys*. 72, 1748 (1992).
28. D. M. Gruen, A. R. Krauss, C. D. Zuiker, R. Csencsits, L. J. Terminello, J. A. Carlisle, I. Jimenez, D. G. J. Sutherland, D. K. Shuh, W. Tong, and F. J. Himpsel, *Appl. Phys. Lett.* 68, 1640 (1996).
29. R. J. Nemanich, J. T. Glass, G. Lucovsky, and R. E. Shroder, *J. Vac. Sci. Technol. A* 6, 1783 (1988).
30. A. R. Krauss, D. M. Gruen, D. Zhou, T. G. McCauley, L. C. Qin, T. Corrigan, O. Auciello, and R. P. H. Chang, *Mater. Res. Soc. Symp. Proc.* 509 (1998).
31. P. Keblinski, S. R. Phillpot, D. Wolf, and H. Gleiter, *Nanostruct. Mater.* 12, 339 (1999).
32. M. P. Siegal, D. R. Tallant, P. N. Provencio, D. L. Overmyer, R. L. Simpson, and L. J. Martinez-Miranda, *Appl. Phys. Lett.* 76, 3052 (2000).
33. M. Zarrabian, N. Fourches-Coulon, G. Turban, C. Marhic, and M. Lancin, *Appl. Phys. Lett.* 70, 2535 (1997).
34. S. Ravi, P. Silva, S. Xu, X. Tay, H. S. Tan, and W. I. Milne, *Appl. Phys. Lett.* 69, 491 (1996).
35. J. E. Field, "The Properties of Diamond." Academic, London, 1979.
36. A. Badzian, T. Badzian, and S. Lee, *Appl. Phys. Lett.* 62, 3432 (1993).
37. R. Locher, C. Wild, N. Herres, D. Behr, and P. Koidl, *Appl. Phys. Lett.* 65, 34 (1994).

38. S. Jin and T. D. Moustakas, *Appl. Phys. Lett.* 65, 403 (1994).
39. D. Zhou, A. R. Krauss, L. C. Qin, T. G. McCauley, D. M. Gruen, T. D. Corrigan, R. P. H. Chang, and H. Gnaser, *J. Appl. Phys.* 82, 4546 (1997).
40. S. J. Harris, *Appl. Phys. Lett.* 56 2298, (1990).
41. S. Bohr, R. Haubner, and B. Lux, *Appl. Phys. Lett.* 68, 1075 (1996).
42. R. Corvin, J. G. Harrison, S. A. Catledge, and Y. K. Vohra, *Appl. Phys. Lett.* 80, 2550 (2002).
43. R. Samlenski, C. Haug, R. Brenn, C. Wild, R. Locher, and P. Koidl, *Appl. Phys. Lett.* 67, 2798 (1995).
44. W. Muller-Sebert, E. Worner, F. Fuchs, C. Wild, and P. Koidl, *Appl. Phys. Lett.* 68, 759 (1996).
45. G. Z. Cao, J. J. Schermer, W. J. P. Van Enckevort, W. A. L. M. Elst, and L. J. Giling, *J. Appl. Phys.* 79, 1357 (1996).
46. D. M. Gruen, S. Liu, A. R. Krauss, J. Luo, and X. Pan, *Appl. Phys. Lett.* 64, 1502 (1994).
47. C. Zuiker, A. R. Krauss, D. M. Gruen, X. Pan, J. C. Li, R. Csencsits, A. Erdemir, C. Bindal, and G. Fenske, *Thin Solid Films* 270, 154 (1995).
48. A. Erdemir, C. Bindal, G. R. Fenske, C. Zuiker, R. Csencsits, A. R. Krauss, and D. M. Gruen, *Diam. Films Tech.* 6, 31 (1996).
49. D. M. Gruen, C. D. Zuiker, A. R. Krauss, and X. Pan, *J. Vac. Sci. Technol. A* 13, 1628 (1995).
50. D. Zhou, D. M. Gruen, L. C. Qin, T. G. McCauley, and A. R. Krauss, *J. Appl. Phys.* 84, 1981 (1998).
51. D. Zhou, T. G. McCauley, L. C. Qin, A. R. Kruass, and D. M. Gruen, *J. Appl. Phys.* 83, 540 (1998).
52. S. Bhattacharyya, O. Auciello, J. Birell, J. A. Carlisle, L. A. Curtiss, A. N. Goyette, D. M. Gruen, A. R. Krauss, J. Schlueter, A. Sumant, and P. Zapol, *Appl. Phys. Lett.* 79, 1441 (2001).
53. A. N. Goyette, J. E. Lawler, L. W. Anderson, D. M. Gruen, T. G. McCauley, D. Zhou, and A. R. Krauss, *Plasma Sources Sci. Technol.* 7, 149 (1998); *J. Phys. D* 31, 1975 (1998).
54. D. M. Gruen, *MRS Bull.* 23, 32 (1998).
55. Y. K. Vohra and S. A. Catledge, U. S. Patent 6, 183, 818, 2001.
56. S. A. Catledge and Y. K. Vohra, *Mater. Res. Soc. Symp. Proc.* 555 (1999).
57. S. A. Catledge, in Microstructure and Mechanical Properties of Diamond Films Grown on Ti-6Al-4V, Dissertation, University of Alabama at Birmingham, 1999.
58. T. Sharda, M. Umeno, T. Soga, and T. Jimbo, *Appl. Phys. Lett.* 77, 4304 (2000).
59. H. Hahn, P. Mondal, and K. A. Padmanabhan, *Nanostruct. Mater.* 9, 603 (1997).
60. R. W. Siegel, *Mater. Sci. Forum* 851, 235 (1997).
61. J. E. Carsley, J. Ning, W. W. Milligan, S. A. Hackney, and E. C. Aifantis, *Nanostruct. Mater.* 5, 441 (1995).
62. J. W. Ager and M. D. Drory, *Phys. Rev. B* 48, 2601 (1993).
63. J. Sizemore, R. J. Holfelder, J. J. Vlassak, and W. D. Nix, *Mater. Res. Soc. Symp. Proc.* 383 (1995).
64. R. Ramesham, R. F. Askew, and M. F. Rose, *Mater. Res. Soc. Symp. Proc.* 383 (1995).
65. Q. H. Fan, A. Fernandes, E. Pereira, and J. Gracio, *J. Appl. Phys.* 83, 5588 (1998).
66. M. D. Drory and J. W. Hutchinson, *Mater. Res. Soc. Symp. Proc.* 383 (1995).
67. A. Erdemir, M. Halter, G. R. Fenske, C. Zuiker, R. Csencsits, A. R. Krauss, and D. M. Gruen, *Trib. Trans.* 40, 667 (1997).
68. B. Hall and P. Sheridan, "NIH Panel Makes Recommendations for Temporomandibular Disorders," http://www.hhs.gov/news/press/1996pres/960501.html.
69. D. W. Feigal, "TMJ Implants—A Consumer Informational Update," http://www.fda.gov/cdrh/consumer/tmjupdate.html.
70. TMD: Temporomandibular Disorders, NIH Publication No. 94-3487, Natl. Inst. of Health, Bethesda, MD, 1994.
71. T. Gillespi III and M. Richardson, "TMJ Anatomy," http://www.cpmcnet.columbia.edu/dept/dental/Dental_Educational_Software/TMJ/TMJAnatomy.html.
72. D. Christenson, *TMJ Sci.* 1, 9 (2001).
73. C. R. Wheeless, "Wheeless' Textbook of Orthopaedics," http://ortho-u.net/.
74. M. J. Jasty, W. E. Floyd, A. L. Schiller, S. R. Goldring, and W. H. Harris, *J. Bone Joint Surg. Am.* 68, 912 (1986).
75. R. Lappalainen, H. Heinonen, A. Anttila, and S. Santavirta, *Diam. Relat. Mater.* 7, 482 (1998).
76. A. V. Lombardi, T. H. Mallory, B. K. Vaughn, and P. Drouillard, *J. Bone Joint Surg. Am.* 71, 1337 (1989).
77. H. J. Agins, N. W. Alcock, M. Bansal, E. A. Salvati, P. D. Wilson, P. M. Pellici, and P. G. Bullogh, *J. Bone Joint Surg. Am.* 70, 347 (1988).
78. R. Venugopalan, M. A. George, J. J. Weimer, and L. C. Lucas, *Biomaterials* 20, 1709 (1999).
79. C. Machado, M. C. Castaneda, J. Diaz, P. Pineros, and L. Suarez, unpublished data.
80. J. Lemons, P. J. Louis, and P. Beck, unpublished data.
81. G. S. Leventhal, *J. Bone Joint Surg.* 33A, 473 (1951).
82. D. M. Brunette, "Titanium in Medicine." Springer, Berlin, 2001.
83. P. I. Branemark, *J. Prosthet. Dent.* 50, 399 (1983).
84. ASTM F86-91: Standard practice for surface preparation and marking of metallic surgical implants, in "Annual Book of ASTM Standards: Medical Devices and Services," Vol. 10.01, pp. 6–8. American Society for Testing and Materials, Philadelphia, 1995.
85. M. Iwaki, S. Sato, K. Takahashi, and H. Sakairi, *Nucl. Instrum. Methods* 210, 1129 (1983).
86. Y. Pleskov, A. Sakharova, M. Krotova, L. Bouilov, and B. Spitsyn, *J. Electroanal. Chem.* 228, 18 (1987).
87. R. Ramesham and M. Rose, *Thin Solid Films* 300, 144 (1996).
88. J. Xu, M. Granger, Q. Chen, J. Strojek, T. Lister, and G. Swain, *Anal. Chem.* 97, 591A (1997).
89. R. Schnupp, R. Kühnhold, G. Temmel, E. Burte, and H. Ryssel, *Biosensors Bioelectron.* 13, 889 (1998).
90. G. Swain, A. Anderson, and J. Angus, *MRS Bull.* 23, 56 (1998).
91. J. Moon, S. Park, Y. Lee, G. Bang, Y. Hong, C. Park, and I. Jeon, *J. Electroanal. Chem.* 464, 230 (1999).
92. F. Beck, W. Kaiser, and H. Krohn, *Electrochim. Acta* 45, 4691 (2000).
93. B. Fausett, M. Granger, M. Hupert, J. Wang, G. Swain, and D. Gruen, *Electroanalysis* 12, 7 (2000).
94. D. Dowling, P. Kola, K. Donnelly, T. Kelly, K. Brumitt, L. Lloyd, R. Eloy, M. Therin, and N. Weill, *Diamond Relat. Mater.* 6, 390 (1997).
95. S. Mitura, A. Mitura, P. Niedzielski, and P. Couvrat, *Chaos Solitons Fractals* 10, 2165 (1999).
96. P. Kovacs, *Trans. Soc. Biomater.* 19, 272 (1993).
97. R. J. Solar, S. R. Pollack, and E. Korostoff, *J. Biomed. Mater. Res.* 13, 217 (1979).
98. M. Pourbaix, *Biomaterials* 5, 122 (1984).
99. M. Fries, R. Venugopalan, and Y. Vohra, unpublished data.
100. R. Venugopalan and J. Gaydon, Corrosion Behavior of Common Implant Alloys, Technical Note, EG&G, Princeton Applied Research, Oak Ridge, TN, 1998.
101. S. Woodard, A Novel Metalloceramic Coating for Reduced UHMWPE Wear in Biomedical Devices, Master's Thesis, University of Alabama at Birmingham, 2002.
102. A. Sakharova, L. Nyikos, and Y. Pleskov, *Electrochim. Acta* 37, 973 (1992).
103. J. van de Lagemaat, G. Z. Cao, G. Janssen, D. Vanmaekelbergh, and J. J. Kelly, *Electrochem. Soc. Proc.* 95, 103 (1995).
104. M. Ogiso, N. Nagabayashi, T. Matsumoto, M. Yamamura, and R. R. Lee, *J. Biomed. Mater. Res*. 30, 109 (1996).
105. W. R. Lacefield, *Ann. NY Acad. Sci.* 72 (1988).

106. W. R. Lacefield, E. D. Rigney, L. C. Lucas, J. Ong, and J. B. Gantenberg, in "Ceramics in Substitutive and Reconstructive Surgery," pp. 167–176. Elsevier, New York, 1991.
107. J. L. Ong, L. C. Lucas, W. R. Lacefield, and E. D. Rigney, *Biomaterials* 13, 249 (1992).
108. R. R. Kumar and M. Wang, *Mater. Sci. Eng. A* 338, 230 (2002).
109. R. R. Kumar and M. Wang, *Mater. Lett.* 49, 15 (2001).
110. S. W. K. Kweh, K. A. Khor, and P. Cheang, *Biomaterials* 21, 1223 (2000).
111. G. Muralithran and S. Ramesh, *Ceram. Int.* 26, 221 (2000).
112. C. Zhang, Y. Leng, and X. Zhang, *J. Biomed. Mater. Res.* 50, 267 (2000).
113. C. Zhang, Y. Leng, and J. Chen, *J. Biomed. Mater. Res.* 56, 342 (2001).
114. S. A. Catledge, Y. K. Vohra, and W. R. Lacefield, unpublished data.
115. J. L. Ong, L. A. Harris, L. C. Lucas, W. R. Lacefield, E. D. Rigney, *J. Am. Ceram. Soc.* 74, 230 (1991).
116. R. Y. Whitehead, L. C. Lucas, and W. R. Lacefield, *Clin. Mater.* 12, 31 (1993).
117. W. R. Lacefield, L. C. Lucas, E. D. Rigney, J. L. Ong, T. S. Chen, and J. B. Gantenberg, in "Surface Modification Technologies VI" (T. S. Sudarshan and J. F. Braza, Eds.), pp. 591–603. The Minerals, Metals and Materials Society, 1993.
118. J. L. Ong, L. C. Lucas, W. R. Lacefield, and E. D. Rigney, *Biomaterials* 13, 249 (1992).
119. R. Y. Whitehead, W. R. Lacefield, and L. C. Lucas, *J. Biomed. Mater. Res.* 27, 1501 (1993).
120. T. S. Chen and W. R. Lacefield, *J. Mater. Res.* 9, 565 (1994).
121. L. C. Lucas, W. R. Lacefield, J. L. Ong, and R. Y. Whitehead, *Colloids Surfaces* 77, 141 (1993).
122. W. R. Lacefield, *J. Oral Implantol.* 20, 214 (1994).
123. H. Dasarathy, C. Riley, H. Coble, W. R. Lacefield, and G. Maybee, *J. Biomed. Mater. Res.* 31, 81 (1996).
124. S. Wang, W. R. Lacefield, and J. E. Lemons, *Biomaterials* 17, 1965 (1996).
125. W. R. Lacefield, *Implant Dentistry* 7 (1998).
126. H. Zeng, K. K. Chittur, and W. R. Lacefield, *Biomaterials* 20, 377 (1999).
127. H. Zeng, K. K. Chittur, and W. R. Lacefield, *Biomaterials* 20, 443 (1999).
128. H. C. Amstutz, P. Campbell, N. Kossovsky, and I. Clarke, *Clin. Orthopedics* 276, 7 (1992).
129. S. R. Goldring, S. Schiller, M. Roelke, M. Rourke, S. A. O'Neill, and W. H. Harris, *J. Joint Bone Surgery A* 65, 575 (1983).
130. J. H. Dubleton, *J. Biomater. Appl.* 3, 3 (1988).
131. B. Weightman and D. Light, *Biomaterials* 7, 20 (1986).
132. G. Derbyshire, J. Fisher, D. Dowson, C. S. Hardaker, and K. Brummitt, *Wear* 181–183, 258 (1995).
133. J. R. Cooper, D. Dowson, and J. Fisher, *Clin. Mater.* 14, 295 (1993).
134. J. M. Cuckler, J. Bearcroft, and C. M. Asgian, *Clin. Orthopedics Relat. Res.* 317, 57 (1995).
135. R. A. Buchanan, E. D. Rigney, and J. M. Williams, *J. Biomed. Mater. Res.* 21, 355 (1987).
136. T. B. Rostlund, B. Albrektsson, T. Albrektsson, and H. McKellop, *Biomaterials* 10, 176 (1989).
137. A. Wisbey, P. J. Gregson, and M. Tuke, *Biomaterials* 8, 477 (1987).
138. T. Burakowski and T. Wierzchon, *Surface Eng. Metals* 176 (1999).
139. M. I. Jones, I. R. McColl, D. M. Grant, K. G. Parker, and T. L. Parker, *J. Biomed. Mater. Res.* 52, 413 (2000).
140. R. Hubler, A. Schroer, W. Ensinger, G. K. Wolf, W. H. Schreiner, and I. J. R. Baumvol, *Surf. Coat. Tech.* 60, 561 (1993).
141. M. Benmalek, P. Gimenez, J. P. Peyre, and C. Tournier, *Surf. Coat. Technol.* 48, 181 (1991).
142. A. A. Voevodin, S. V. Prasad, and J. S. Zabinski, *J. Appl. Phys.* 82, 855 (1997).
143. A. A. Voevodin, M. S. Donley, J. S. Zabinski, and J. E. Bultman, *J. Appl. Phys.* 77, 534 (1995).
144. A. A. Voevodin, S. D. Walck, J. S. Solomon, P. J. John, D. C. Ingram, J. S. Zabinski, and M. S. Donley, *J. Vac. Sci. Technol. A* 14, 1927 (1996).
145. M. J. Pappas, G. Makris, and F. F. Buechel, *Clin. Orthopedics Relat. Res.* 317, 64 (1995).
146. C. D. Peterson, B. M. Hillberry, and D. A. Heck, *J. Biomed. Mater. Res.* 22, 887 (1988).
147. R. A. Kant and B. D. Sartwell, *Mater. Sci. Eng.* 90, 357 (1987).
148. S. Catledge et al., *J. Nanosci. Nanotechnol.* 2, 293 (2002).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanostructured Zeolite Films

Yushan Yan, Huanting Wang

University of California, Riverside, California, USA

CONTENTS

1. INTRODUCTION

1.1. Zeolites

Zeolites and zeolite-type materials (hereafter referred to as zeolites for simplicity) are a class of crystalline oxides with uniform channels and/or cages ranging from 0.3 to 2 nanometers (nm). Structurally, a zeolite consists of three dimensionally linked tetrahedra and each tetrahedron has one tetrahedrally coordinated atom (hereafter referred to as T-atom) (e.g., Si, Al, P, B, Ga, Ge, Zn, Be, etc.) at the center and four oxygen atoms at the corners. Every corner oxygen in the zeolite structure is shared by two neighboring tetrahedra [1, 2]. Depending on the T-atom, the zeolite framework could be charged or neutral. When the framework is charged, balancing ions are needed, and this is the origin of ion exchange capability of zeolites. According to their structural symmetry and topology, zeolites are classified into different framework types. More than 130 framework types with numerous compositional variations are known [2, 3]. Each framework type is assigned with a unique three-letter code by the International Zeolite Association (www.izc-online.org). For example, zeolite ZSM-5 is designated as MFI. The three-letter codes will be followed whenever appropriate in this review.

At present, zeolite materials are commercially used as shape/size selective catalysts in petrochemical industry and molecular sieving separation media for fixed gases and hydrocarbons. For both applications, zeolites are used in powder composite form such as pellets and granules. Zeolite catalysis has been an important and expanding field and several recent reviews are available [4–6]. Zeolites are good acid catalysts. Zeolites have also found use in base and oxidation catalysis. Immobilization of catalytically active guests such as ship-in-the-bottle complexes, chiral guests, and enzymes has also been studied [4]. One key benefit of zeolite catalysts is their shape-selectivity. For example, zeolites are good catalysts for conversion of two-ring and three-ring polycyclic hydrocarbons into value-added intermediates for making specialty chemicals and monomers of new polymer materials such as advanced engineering plastics, polyester fibers and films, and liquid crystal polymers. New developments in this area were discussed in a recent review covering reactions over large-pore zeolite catalysts [5]. Another recent review focused on heterogeneous chiral catalysis and the use of zeolites containing chelate complexes was discussed [6].

Separation and purification of gases or liquids by selective adsorption into zeolite adsorbents are well established commercially. For example, air separation by pressure swing adsorption (PSA) has been practiced for many years. Zeolites have also been used for separation of xylene isomers. Another important use of zeolites is as detergent builders. Little discussion of this application has appeared in the open literature although numerous patents have been pursued or issued [7–14]. The use of zeolite LTA as detergent builders is very well documented [7–10].

1.2. Related Recent Reviews

Several reviews related to zeolite films and their synthesis and applications have been published [15–25]. This section is intended to provide a brief analysis of emphases of these reviews. Applications of zeolite films as membranes and catalytic coatings have been discussed in several short reviews [15–17]. In 1996, Bein published a comprehensive review on zeolite membrane and layers [18]. This review had very

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (763–781)

extensive coverage on sensor applications and strategies for fabrication of oriented crystal layers. In 1999, Tavolaro and Drioli published a review on zeolite membranes which provided a comprehensive list in convenient table format of various zeolite membranes (zeolite and support) and their separation performances [20]. Polymer-zeolite composite membranes were also included. In 2000, Caro et al. published a review on zeolite membranes which provided critical analyses on important issues such as how crystal orientation and grain boundaries in a zeolite membrane affect its thermal stability and diffusion resistance [23]. In 2000, Tsapatsis et al. published a review that focused on the relationship between microstructure and permeability of MFI membranes [21]. This review had extensive discussion of the three hydrothermal synthesis methods, especially seeded growth. Another two reviews were published recently [24, 25]. Although these two did not focus on zeolite films, they, however, provided discussions on new applications of zeolite materials and many of these new applications require zeolite to be in a film form.

All currently available reviews on polycrystalline zeolite films focused on membrane and membrane reactor applications. The goal of this review is to take an integrated approach that includes zeolite films with different configurations (e.g., polycrystalline, matrix composite, and layers), provide an introduction and analyses of film preparation methods and film applications, and provide a comprehensive list of zeolite films published up to 2002 in a format that can be easily referenced and expanded.

1.3. Structure of This Review

This review begins with a brief introduction of zeolite materials and their structures. Current commercial applications of zeolites are briefly discussed (Section 1.1). Recent reviews related to zeolite films are analyzed (Section 1.2). Section 2 provides a brief discussion of types of zeolite-based films followed by Section 3 on demonstrated synthesis methods. Section 4 documents applications for zeolite films, including membrane, membrane reactor, adsorption, catalysis, sensor, low-k dielectrics, corrosion-resistant coating, hydrophilic coating, heat pump, and thermoelectrics. To keep Sections 3 and 4 brief, only significant examples that are most familiar to the authors were selected to illustrate the concepts. A comprehensive collection of all reported zeolite films are provided in Section 5. Most of the films covered in Section 5 are not discussed in previous sections and are organized according to their three-letter codes (except MFI, FAU, and LTA) for easy reference. Some recent developments are discussed in Section 6. Concluding remarks and future perspectives are provided in Section 7 followed by acknowledgments, and references.

2. TYPES OF NANOSTRUCTURED ZEOLITE FILMS

2.1. Polycrystalline Zeolite Film

Polycrystalline zeolite films consist of intergrown zeolite crystals with no apparent significant cracks or pinholes (e.g., [26]) (Fig. 1). These films are made of zeolite only, without any nonzeolite components (e.g., amorphous silica or polymer). They are normally supported on a substrate although free-standing films have also been synthesized [27]. Polycrystalline films are primarily prepared by hydrothermal synthesis methods including *in-situ* crystallization [26], seeded growth [28], and vapor transport [29] (see Section 3 for details) and have potential use in all of the applications discussed in this review (see Section 4 for details).

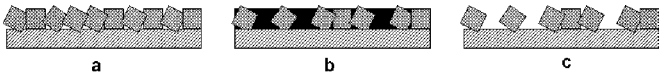


Figure 1. Schematic of the three types of nanostructured zeolite films. (a) a polycrystalline zeolite film, (b) a zeolite matrix composite film, and (c) a zeolite crystal layer.

2.2. Zeolite Matrix Composite Film

A zeolite matrix composite film is defined as a film in which isolated zeolite crystals are imbedded in a solid matrix that is either inorganic (e.g., silica and carbon) [30, 31, 32] organic (e.g., polymer [33]) (Fig. 1). For an inorganic matrix, these films are often supported while for an organic polymer matrix they normally are free-standing. The selection between an inorganic and an organic phase as the imbedding matrix depends strongly on the intended application. For example, a silica matrix would be ideal for applications where thermal stability is required [30], whereas a polymer matrix is a natural choice when flexibility is desired [33]. These composite films are usually prepared by wash-coating, sol–gel processing, and polymer film casting (see Section 3 for details).

2.3. Zeolite Crystal Layer

Another form of zeolite films is zeolite crystal layer which consists of isolated crystals deposited on a solid substrate [18] (Fig. 1). The substrate could be a variety of materials such as metal, ceramic, or silicon wafer. Although polycrystalline and matrix composite films could be prepared as a free-standing film, crystal layers have to be supported. There has been exciting fundamental research carried out in this area [34, 35]. However, demonstrated applications have been limited to sensors. Organic linker approach appears very promising for preparation of this type of film [36–39].

3. DEMONSTRATED SYNTHESIS STRATEGIES OF ZEOLITE FILMS

3.1. Hydrothermal Synthesis

Zeolites are normally synthesized hydrothermally in a sealed reactor under autogeneous pressure. The synthesis temperature is normally below 200 °C and the pressure inside the reactor is about several atmospheric pressures depending on synthesis temperature and the chemical composition of the synthesis mixture. The term "hydrothermal" by definition means that water is the solvent of choice in the synthesis. Zeolite synthesis has also been successfully carried out in nonaqueous solvent. When a nonaqueous solvent is

used, the process is sometimes called solvothermal synthesis. So far, all known syntheses of zeolite films used hydrothermal synthesis although solvothermal process may offer distinctive advantages. A common feature of the hydrothermal methods is that they produce pure polycrystalline "binderless" zeolite films while "binders" are needed in the nonhydrothermal techniques.

3.1.1. *In-situ* Crystallization

In-situ crystallization refers to a film deposition process where the substrate to be coated is directly submerged in the synthesis mixture and the zeolite crystals are formed "*in-situ*" on the substrate surface [26, 40]. The term "*in-situ* crystallization" originates from the fact that no preformed zeolite crystals are used in the synthesis. *In-situ* crystallization could be considered as a natural extension of the normal hydrothermal synthesis of zeolite powders. However, it is important to recognize that synthesis of a continuous zeolite film on a substrate could be fundamentally different from that of loose zeolite powder [26] (Fig. 2). To form high quality zeolite films, the synthesis composition and crystallization conditions have to be conducive to good intergrowth between the zeolite crystals and good adhesion of the zeolite film to the substrate. This requirement makes composition optimization very often essential [26, 41]. Many of the unsuccessful preparations for a functional zeolite membrane in the early days could be attributed to direct adaptation of a documented recipe for powder synthesis.

The *in-situ* crystallization process has several important advantages over other deposition processes. It is a simple one-step process while all other methods that will be presented in this review involve multisteps. Furthermore, *in-situ* crystallization has been clearly demonstrated to be able to coat surfaces of complex shape and in confined spaces [42]. This capability is important and very often essential for many practical applications. Corrosion-resistant coatings [42] and hydrophilic coatings on heat exchangers [43] are clearly such examples. For adsorption, catalysis, and heat pump applications, zeolite coatings on a structured monolithic support are needed. Many of these monolithic supports possess tiny and sometimes wavy channels whose internal surface has to be coated. Wash-coating, sol–gel processing, assembly with organic linkers, seeded-growth, and the vapor-phase transport method either have clear significant difficulties in coating these surfaces or have not been clearly demonstrated.

The most favorable situation for *in-situ* crystallization is that the film formation proceeds through a heterogeneous nucleation process that produces a dense layer of nuclei/gel on the surface and the bulk crystallization of loose crystals is minimized or avoided. Minimization of the production of zeolite crystals in the bulk avoids clogging of small channels in a structured support such as a heat exchanger. This is clearly the case for some of the ZSM-5 (MFI) films and this could be the major reason that they have been successfully demonstrated for a number of applications including membranes [26], corrosion-resistant coatings [42], and hydrophilic coatings [43]. The presence of a small amount of aluminum in the synthesis mixture is essential for the heterogeneous mechanism to dominate because the formation of pure-silica MFI film appears to proceed via a significantly different mechanism [41, 44]. However, a heterogeneous nucleation/growth process and a suitable synthesis composition may not be available for all zeolites and this could be a limitation for *in-situ* crystallization.

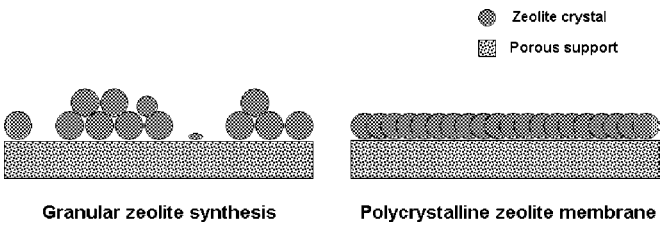


Figure 2. Schematic illustration of synthesis of normal zeolite granules versus a polycrystalline zeolite film.

3.1.2. Seeded Growth

Seeded growth refers to a film deposition process where a layer of preformed zeolite crystals is deposited on the substrate prior to its subsequent hydrothermal treatment [28]. Several methods for introducing the seed layer onto the substrate have been developed and these include rubbing the substrate surface with zeolite crystals [45], laser ablation of zeolite powders [46], and colloidal zeolite particle seeding through dip-coating [28]. The advantages and disadvantages of each seeding method have been thoroughly discussed in a recent review [21]. Colloidal particle seeding appears to be a frequent choice for seeded growth method.

The main advantage of seeded growth is the possibility of decoupling of nucleation and crystal growth. Since the film formation begins with a dense layer of zeolite colloidal particles, and if an appropriate synthesis mixture is developed for the secondary growth, secondary nucleation can be minimized or avoided. Another advantage of the seeded growth is that the growth process can be carried out at a temperature lower than *in-situ* crystallization if desired, although the deposition rate can be low.

The limitation of the seeded growth method is the need to synthesize small dispersible zeolite particles. This often requires sophisticated synthesis engineering, and at present appropriate synthesis composition and crystallization conditions are still not available for some zeolites [21]. In principle, zeolite particles with a small diameter are preferred for seeded growth. However, the yield of zeolite nanoparticles is often low for small diameter particles (e.g., 50 nm) [47]. Although possible in principle, it has not been clearly demonstrated that seeded growth can coat surfaces of complex shape and in confined spaces.

3.1.3. Vapor Phase Transport

Vapor phase transport refers to a film deposition process where a reactive gel layer is deposited on the substrate prior to its subsequent hydrothermal/solvothermal treatment [29]. The vapor introduced in the second step could be the organic structure-directing agent (SDA) or water depending on the zeolite system. Such an approach has been used to prepare supported or free-standing zeolite films. Since SDA molecules are either already concentrated in the gel layer or only needed to treat a minimized amount of gel on the

substrate, vapor phase transport has the benefit of minimized use of SDA, which is often the cost driver for zeolite synthesis.

Similar to seeded growth, vapor phase transport method is a two-step process. For vapor phase transport to work, a reactive gel has to be prepared. So far, a reactive gel is still not available for some zeolites. Another major concern for vapor phase transport is that the gel density is usually much lower than that of zeolite crystals, and thus the crystallization of the gel layer tends to produce intercrystal voids. Consequently, to produce a continuous film, the voids must be filled through the dissolution-recrystallization process.

3.2. Nonhydrothermal Synthesis

One common feature of the nonhydrothermal techniques is that they all use preformed zeolite crystals. In addition, they all use some form of "binder."

3.2.1. Wash-Coating, Sol–Gel Processing, and Polymer Film Casting

Wash-coating is a well-established industrial process for deposition of high-surface-area catalyst support on a structured monolith (e.g., ceramic cordierite honeycomb). Manufacturing of automotive three-way catalysts involves the use of wash-coating of γ-alumina. Typically, fine alumina particles that could be produced by ball milling are blended with an inorganic binder to form a slurry. The slurry is then flown through the channels of the monolith to leave behind a layer of slurry on the channel surface. Air-knifing is usually employed to blow away the extra slurry before the coated monolith is dried and then sintered to establish binding within the film and between the film and the monolith support. This method can be extended straightforwardly to zeolite coatings with silica or alumina-based binder [48]. The key benefit of the wash-coating method is that it is a well-established method and a large amount of know-how is available. However, it has several drawbacks. First, a binder has to be used and this occupies valuable space in the film and also tends to clog zeolite pores at high concentration. Second, high-temperature sintering is needed for binding and this could be incompatible with metallic substrates. Third, this method also has difficulty in handling very small and nonstraight channels (e.g., wavy channels).

Sol–gel processing is also well established [49]. After blending zeolite crystals (usually micrometer-sized) in an appropriate sol (e.g., hydrolyzed tetraethylorthosilicate (TEOS)), a standard sol–gel film formation process such as spin-coating and dip-coating can be used. For this strategy to work, zeolite crystals have to be synthesized first. One simple, but significant extension of this technique to zeolite nanocrystals, was recently developed [30, 31]. Zeolite nanocrystals are usually synthesized hydrothermally and the synthesis solution contains both nanocrystals and amorphous silica components. To retrieve the nanocrystals, a long period of high-speed centrifugation has to be performed. If a sol–gel matrix composite film is the desired product, however, the film can be simply formed by spin-coating or dip-coating by using the nascent synthesis solution without going through centrifugation separation. This has proven to be very promising for preparation of zeolite low-k dielectric films. Another recent study of zeolite composite films has shown that they could be used for photovoltaics [50].

Polymer film casting again is a well-established industrial process. It can produce polymeric membranes for gas and liquid separations. Once zeolite particles are incorporated into the polymer solution, this process could be readily extended for the preparation of polymer-zeolite mixed matrix membranes [51–54]. The main motivation for this composite approach is to combine high selectivity of the zeolite and the excellent processability of the polymer into one membrane.

3.2.2. Assembly of Micron-Sized Crystals Through Organic Linkers

A newly developed method for depositing zeolite layers is the assembly of micrometer-sized zeolite crystals using organic linkers [36, 39, 55–57]. In this method, the substrate and the zeolite crystals are both functionalized in such a way that a covalent bond could be formed between them but not among the zeolite crystals. This favors the formation of a monocrystal layer on the substrate [58]. A wide variety of well-known chemistry could be exploited in this method [59]. This strategy can be used for different zeolites and different substrates [38]. Patterning can also be readily achieved with this synthesis strategy [37, 60]. However, this method produces isolated crystal layers whose application may be limited. The adhesion of the crystals to the substrate is also weak. Nonetheless, this layer could be used as a seed layer for subsequent hydrothermal treatment.

4. APPLICATIONS OF ZEOLITE FILMS

4.1. Membrane

Zeolite membranes can be used for separations of gases, vapors, and liquids. This section focuses on polycrystalline zeolite films and polymer-zeolite mixed matrix films, which are also useful as membranes, and will be briefly discussed in the section of recent developments. Single crystal membranes have been studied. They are important tools for the fundamental study of mass transport through membranes, but are not suitable for large-scale applications [61]. For polycrystalline zeolite membranes, these films are required to be continuous and free from intercrystal porosity. They are also normally deposited on porous substrates for mechanical stability [62–66], although free-standing zeolite membranes have also been synthesized [67, 68]. Because of their extremely uniform pore size, zeolite membranes could have exceptionally high selectivity. Furthermore, due to their inorganic nature, zeolite membranes are more stable than polymeric membranes at high temperatures and in aggressive chemical environments. The high thermal and chemical stability make zeolite membranes suitable for membrane reactor applications. A large fraction of the reports on zeolite membranes are on MFI (ZSM-5 or silicalite) membranes because MFI membranes are relatively easy to prepare and have important practical use. Thus, MFI membranes are used exclusively in this section to illustrate the important developments of zeolite membranes. However, many of the techniques developed for MFI membrane are directly applicable to other zeolite membranes.

MFI membranes have been commonly prepared on porous alumina and stainless steel supports. The butane isomers have slightly different kinetic diameters (n-butane = 4.3 Å, i-butane = 5.0 Å) and both diameters are very close to the pore size of MFI (5.5 Å). Thus n-butane/i-butane selectivity is often used as a measure of the quality of MFI membranes. A high n-butane/i-butane selectivity indicates a MFI membrane of high quality (Fig. 3). Using *in-situ* crystallization, polycrystalline MFI membranes prepared on α-alumina showed n-butane/i-butane selectivity of 31 at 185 °C [26]. Seeded growth [69] was also successfully used for MFI membranes with n-butane/i-butane selectivity of 28–62 at room temperature. Very recently, vapor phase transport method [70] was used to prepare MFI films but permeations results were not reported.

The two key performance parameters for a membrane are permeation rate and selectivity. It was shown that under *in-situ* crystallization conditions for ZSM-5 membranes, there was significant penetration of siliceous material into the pore of the substrate. This penetration was believed to decrease the permeation rate and selectivity. To reduce this penetration, a temporary carbon barrier was introduced into the substrate pores before *in-situ* crystallization (Fig. 4). The carbon barrier was removed by calcination after membrane synthesis. Both flux and selectivity were improved by the use of the temporary carbon barrier [71].

It was also realized that ZSM-5 membranes by *in-situ* crystallization may have micro-defects (i.e., defect diameter < 2 nm). Clearly, if these defects could be removed, the membrane selectivity could be improved. With this goal in mind, a selective coking method was developed and n-butane/i-butane selectivity was improved from 45 to 322 at 185 °C. Triisopropylbenzene (TIPB) was chosen as the coking precursor because it is (kinetic diameter = 8.4 Å) too large to enter the MFI pores (pore diameter = 5.5 Å), so that coking selectively occurs in the microdefects while the MFI pores are intact [72] (Fig. 5). Chemical vapor deposition (CVD) was also used for modification of zeolite membranes to improve shape-selectivity of MFI membranes through counter-diffusion of tetramethylorthosilicate (TMOS)/O_3 [73, 74].

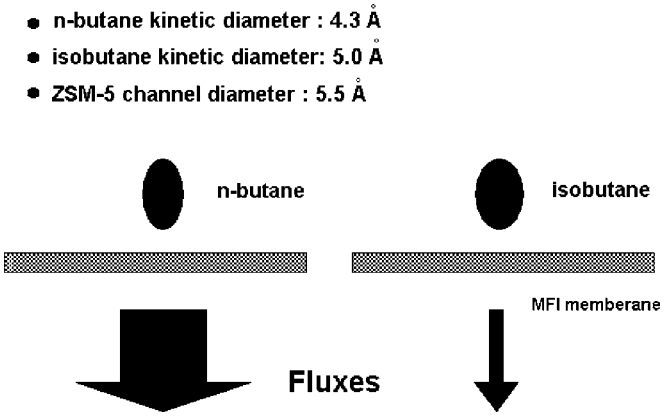


Figure 3. Schematic illustration of molecular sieving effect of a MFI membrane toward butane isomers. Small difference in kinetic diameters results in large difference in fluxes.

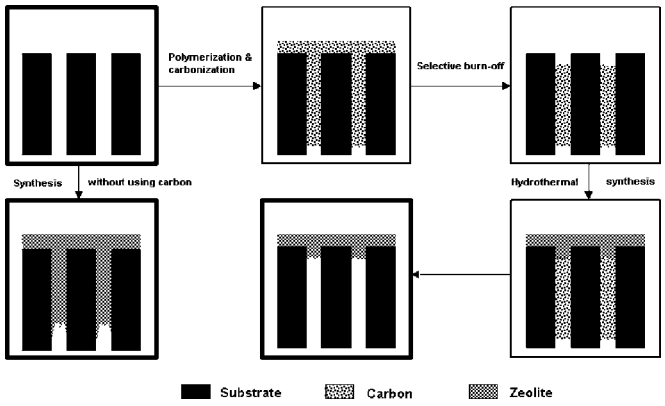


Figure 4. Schematic illustration of the use of a temporary carbon barrier to improve membrane flux.

It is well known that zeolite pore size can be finely modulated by ion exchange [1]. Perhaps the best known example for pore size modulation is zeolite LTA (A). When the cation in LTA (Si/Al ratio of 1) changes from sodium to calcium, the pore size increases from 4 Å to 5 Å. In an attempt to take advantage of this phenomenon, ZSM-5 membranes with Si/Al ratio of 25 were ion exchanged with H^+, Na^+, K^+, Cs^+, Ca^{2+}, and Ba^{2+}. The membranes were stable after exchange and single-gas permeances increased in the order: $K^+ < Ba^{2+} \approx Ca^{2+} < Cs^+ < Na^+ \approx H^+$, which coincided with the decrease in ion size and only the Cs ion did not fit this trend [75]. Another flexibility of a zeolite membrane is isomorphous substitution which could impact the surface hydrophilicity/hydrophobicity. ZSM-5 zeolite membranes with Al, Fe, B, or Ge substituted for Si in the framework were prepared by *in-situ* crystallization onto porous stainless steel supports. Their single-gas and mixture separation properties were compared to silicalite membranes [76].

The synthesis of high-silica zeolite membranes often requires the use of an organic structure directing agent (SDA). These molecules are eventually trapped in the zeolite pores blocking mass transport. Therefore, these molecules have to be removed. Calcination at a high temperature is often used to remove these organic molecules. However, calcination tends to generate cracks in the membrane and reduce selectivity [77]. To understand crack formation, calcination of large SDA-containing MFI crystals was studied [78]. The interesting finding was that there was

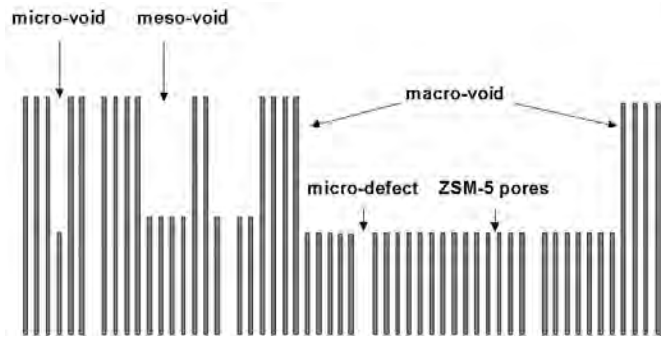


Figure 5. Schematic illustration of a MFI membrane with microdefects that could be selectively filled by coking of triisopropylbezene (TIPB) (after [72]).

positive thermal expansion in a- and c-direction while there was negative expansion in b-direction. This study was later extended to MFI films using *in-situ* X-ray diffraction [79].

To bypass the calcination process, efforts have been made to synthesize MFI membranes without using SDA so that the calcination process is not needed [80]. Hedlund et al. used the seeded growth method. They used SDA-containing crystals as the seed and calcination was performed before the coated substrate was exposed to a synthesis solution that did not contain SDA. The membranes obtained were considered of high quality, but nonzeolitic pathways were evident [81]. *In-situ* crystallization was also used for SDA-free membrane synthesis using a special synthesis composition [82].

To characterize defects, an interesting fluorescence confocal optical microscopy (FCOM) was developed. Fluorescence confocal optical microscopy is a nondestructive tool for the evaluation of zeolite membranes. This technique can provide two-dimensional, optical sections of the membranes with submicron resolution without physical damage to the sample [83]. Another technique for defects characterization is transient permeation which offers much more information than the normal steady state studies [84, 85].

4.2. Membrane Reactor

The high selectivity and high thermal and chemical stability of zeolite membranes make them extremely attractive for membrane reactor applications [86]. The benefits of a membrane reactor are improved selectivity and yield. Selective removal of a product and controlled admission of a reactant are the two commonly used themes. Most membrane reactor applications involve catalysis. The zeolite membrane could actively participate in catalysis. In this configuration, the catalyst is located within the zeolite film [87, 88]. Alternatively, the zeolite membrane could be used passively only as a separation membrane and the catalysts are packed outside the zeolite film [89].

For an active membrane, examples include V-MFI zeolite membrane reactor on alumina tubes for oxidative dehydrogenation of propane. Both MFI and V-MFI membranes were found to produce propene with about 40% selectivity but with higher O_2 and C_3H_8 conversions for the V-MFI [87]. Si-MFI, Ti-MFI, and V-MFI were also prepared by the seeded growth method on porous stainless steel for membrane reactor applications. Transport properties were studied but catalytic performance was not given [88].

For a passive zeolite membrane, the Si-MFI membrane reactor on a porous stainless steel support by the seeded growth method was used for the liquid-phase oligomerization of isobutene. Isooctene was selectively removed from the reaction environment, thus reducing the formation of unwanted C_{12} and C_{16} hydrocarbons. An acid resin catalyst bed was located on the membrane tube side. The membrane reactor produced a very significant increase in the selectivity and yield of isooctenes compared to a conventional fixed bed reactor [89]. zeolite membrane reactors have also been used for selective water removal from gas phase synthesis of methyl-tert-butyl ether (MTBE) from tert-butanol and methanol [90], and for selective hydrogen removal from dehydrogenation of isobutene using Pt-In catalyst [91].

A thermodynamic analysis has also been carried out for CO_2 hydrogenation to methanol in a zeolite membrane reactor where methanol was selectively removed. An increased methanol yield led to reduced consumption of reactant and also allowed the reactor to operate at low pressure and high temperature, which favors the kinetics that reduce the residence time and the reactor volume [92]. ZSM-5 membranes were prepared by seeded growth on porous alumina. After ion-exchanging from Na^+ to the H^+, alkylation of toluene with methanol was carried out and a maximum p-xylene selectivity of 80% was achieved at 450–500 °C [93].

4.3. Sensor

Chemical sensors are important for industrial process control and environmental monitoring. They could also find critical use in medical and defense applications. Two of the critical performance parameters of a sensor are selectivity and sensitivity. It has been shown that when an appropriate zeolite film is used, both of these parameters can be improved [94–96]. A large number of publications [97–124] have appeared in this area, and a comprehensive review on this topic is beyond the scope of this chapter. An in-depth review on zeolite films for chemical sensor applications up to 1996 was provided as a section in a recent review by Bein [18].

The two most commonly used sensor platforms that have been coupled with zeolite films are surface acoustic wave (SAW) and quartz crystal microbalance (QCM) sensors. Using a sol–gel matrix composite film and a SAW sensor platform, Bein et al. formed thin coating of zeolite with pores from 0.4 and 0.8 nm (e.g., zeolite Y, chabazite, and ZSM-5) in hydrolyzed TEOS. High discrimination ratios were observed between molecules entering the zeolite pore system and those excluded from the molecular sieving sensor [94]. In another example, silicalite was grown by *in-situ* crystallization in an oriented way on a quartz substrate to form a well-defined SAW-sensor chemical interface for CO_2 detection. The SAW-sensor device responded well to CO_2, while the sensitivity to water seemed not to be severe [102].

Using an organic linker approach, and a QCM sensor platform, Yan and Bein attached zeolite crystals to the gold electrodes. Thiol-alkoxysilanes on the gold surface served as interfacial layers for the subsequent adhesion of the zeolite crystals to the QCM. The resonance frequency response of zeolite-coated QCM to vapor pulses could be increased up to 500-fold compared to the bare sensor. The regular micropores (0.3–0.75 nm) of the QCM-attached zeolite crystals were found to efficiently control molecular access into the film. Selectivity of the frequency response in excess of 100:1 toward molecules of different size and/or shape could be demonstrated [98]. A seeded growth method has also been used to deposit silicalite film on a QCM sensor for hydrocarbons detection. The seed layer was introduced by electrostatic interaction by modifying the substrate with a cationic polymer. The sensor was evaluated as chemical sensors by exposing it to highly dilute hydrocarbon vapors [109].

4.4. Adsorption

Zeolites in powdery form (e.g., pellets and granules with binders) have been commercially used as effective adsorbents for separation and purification. For example, air

separation can be achieved by using a pressure/temperature swing adsorption process using zeolite adsorbents. The primary motivation for moving from powder to a film configuration is to reduce pressure drop. Other benefits of a film versus packed column include no attrition and potentially better mass transfer and heat management. When hydrothermal synthesis is used, zeolite coating also involves no binder and this could be important because binder tends to clog zeolite pores. Usually for this application, a structured support such as a monolith honeycomb is needed. Literature on this technology dates back to the 1960's when a zeolite coating was used on cryopumping surfaces [125].

In the early 1990's, Davis et al. published a study on the formation of zeolite Y coatings on flat metal surfaces and the use of this kind of coating for adsorption was suggested [126]. The method used was *in-situ* crystallization. It was shown that zeolite crystals were selectively bonded to one side of the metal foil substrate due to gravitational effects during the zeolite crystallization. Dunne and Taqvi et al. [127, 128] have used slurry wash-coating to deposit zeolite coatings on flat metal panels for adsorption applications including moisture and volatile organic compounds (VOC) removal from air. A unique configuration was developed involving a rotating wheel so that adsorption and regeneration can be achieved simultaneously. As mentioned before, the wash-coating deposition process requires a high temperature sintering step (e.g., 650 °C). Monolithic supports have also been used for coatings of ZSM-5 deposited by *in-situ* crystallization and by vapor phase transport [129–131].

4.5. Catalysis

Similar to adsorption, the choice of zeolite coatings on a structured support over a packed bed of zeolite particles (e.g., granules and pellets) is motivated by the desire to lower the pressure drop [132]. Zeolite-coating reactors also could have high catalyst utilization because of facile mass transfer. The use of zeolite coatings dates back to the early 1970's [133, 134]. Wash-coating was used on simple supports such as glass fibers. In a later patent, more complex support was claimed [135]. *In-situ* crystallization that could produce binderless zeolitic coatings is very attractive, and has been used to deposit ZSM-5 coating on monolithic stainless steel structures [132]. Synthesis composition was optimized, and homogeneous zeolitic coatings with a Si/Al ratio of 34 and with coverage of up to 25 g/m^2 were obtained in a single synthesis run of 24 h at 150 to 170 °C. Neither the state of the silicon source nor the curvature of the support surface (monoliths constructed from foil or 35-μm wire gauze) had a significant effect on these figures.

More recent developments on zeolite coatings seem to focus on microreactors [136, 137]. Hydrothermal formation of ZSM-5 zeolitic coatings on stainless steel plates with a microchannel structure has been investigated at different synthesis mixture compositions. The procedures of coating and thermal treatment were optimized. Obtaining a uniform thickness of the coating within 0.5 μm-wide microchannels required a careful control of various synthesis variables. The prepared samples proved to be active in the selective catalytic reduction (SCR) of NO with ammonia. The microreactor showed no mass transfer limitations and a larger SCR reaction rate was observed in comparison with pelletized Ce-ZSM-5 catalysts [136].

4.6. Low-k Dielectrics

Low-k (k for dielectric constant) dielectric materials have been identified as one of the most difficult challenges for interconnects of the future generation integrated circuits (ICs). Many materials have been proposed, studied, or under commercial development as potential candidates for low-k dielectrics. Among these materials, two major classes are dense organic polymers and porous inorganic-based materials. It has been shown that some dense, organic polymers could easily have k between 2 to 3, but there are concerns about their low thermal stability and low heat conductivity. Also, due to their low mechanical strength, polymeric materials may have potential problems with the chemical and mechanical polishing (CMP) process.

Current dielectric material is dense silica that has a $k = 4$. Porous silica has lower k because of the incorporation of air. Sol–gel silica and mesoporous silica have been studied as possible low-k materials. However, these materials have concerns of low mechanical strength and low heat conductivity. Sol–gel silica also has randomly occurring large pores that can cause electrical breakdown. High mechanical strength is needed for the new low-k materials to be compatible with CMP process.

Very recently, zeolites have been studied and demonstrated to be a promising low-k material. Zeolites have very uniform pores and their pore size (<2 nm) is significantly smaller than IC feature size. Therefore, the problem of electrical breakdown should be mitigated. Zeolites have high heat conductivity and mechanical strength due to their dense crystalline structure [30, 138, 139]. Pure-silica MFI films were first developed by using *in-situ* crystallization. These films had a k value of 2.7 and were highly hydrophobic. The elastic modulus was 30–40 GPa by nanoindentation using a film of 0.43 μm thick. This value is much higher than sol–gel silica or mesoporous silica films. Mechanical polishing experiments showed that the film was strongly adhered to the silicon wafer and there was no cracking or delamination during the polishing. Low resistivity, high resistivity, oxide-covered, and nitride-covered wafers have been successfully coated using *in-situ* crystallization method.

Although the films formed by *in-situ* crystallization are strong mechanically, the k value is still high. The solution phase deposition is also of concern to the semiconductor industry. Another process was recently developed by a simple spin-on process of zeolite nanoparticle suspension [30]. The spin-on process is simple and preferred by the semiconductor manufacturers. Therefore, it represents a significant step forward toward the integration of zeolite materials into microelectronic devices. Nanoindentation was performed on a 0.42-μm thick film and the elastic modulus was 16–18 GPa. It is worth mentioning that a modulus of 6 GPa is usually considered to be a threshold value for industrially viable low-k dielectrics.

4.7. Corrosion-Resistant Coating

Aluminum alloys are widely used in aerospace because they are light and mechanically strong. However, some of the aerospace aluminum alloys (e.g., 2024-T3) are very prone

to corrosion. For combating corrosion of aluminum alloys, chromic acid anodization and chromate conversion have thus far proved most effective. Both processes, however, use and release hexavalent chromium, a proven human carcinogen, causing serious environmental and worker health and safety concerns. A chromium-free alternative with equivalent or superior corrosion performance has long been desired. Sol–gel coatings have been shown to be a promising alternative and are cited for their superior adhesion, good barrier properties, and high thermal and chemical stability. However, sol–gel coatings usually require high temperature curing that is not compatible with many metal alloys. It is also difficult to coat surfaces of complex shape and in confined spaces. Electrochemical deposition of conductive polymer coatings is another developing technique for combating corrosion of metal and metal alloys. Thermal stability and adhesion are concerns for pure polymer coatings.

The idea of corrosion-resistant zeolite coatings takes advantage of several well-known facts about zeolites. First, high-silica and pure-silica zeolites are stable thermally and chemically. For example, many high-silica zeolites are thermally stable to 1000 °C, and pure-silica zeolites are stable in all mineral acids except hydrofluoric acid [140]. Thus, the framework of zeolites has the needed corrosion resistance. Second, most of the high-silica zeolite syntheses use SDA molecules, and these molecules are eventually trapped inside the zeolite pores. Normally, for separation and catalysis applications, these molecules have to be removed by high temperature calcination. It was recognized very early on that as-synthesized zeolite membranes could be gastight if they were defect-free. This suggested that as-synthesized zeolite films could be corrosion resistant.

It has been shown recently that as-synthesized SDA-containing high-silica MFI coatings on aluminum alloys have superior corrosion resistance to chromate conversion coatings in strong acids, bases, as well as pitting aggressive media [42]. It was also demonstrated that the *in-situ* crystallization coating deposition process can coat surfaces of complex shape and in confined spaces. In addition, the thermal and mechanical properties of the high-silica MFI coatings were shown to be satisfactory [143]. It was also shown that corrosion resistance is general for all high-silica zeolite coatings such as MFI, BEA, and MTW [141–143], and high-silica MFI coatings can be extended to other aluminum alloys such as 6061-T6.

4.8. Hydrophilic Coating

An environmental control system for a manned spacecraft typically includes a condenser for controlling cabin temperature and humidity. As the moisture-laden air is cooled in the small channels of the condenser, water normally condenses out as water droplets. Under a microgravity environment, these water droplets tend to become entrained in the air stream and carried back into the cabin, failing to achieve water separation. In such an instance, rain could quite literally fall inside the cabin. To effectively remove water from the air, heat exchanger surfaces could be coated with a hydrophilic zeolite coating [43] so that the condensation mode changes from drop-wise to film-wise (Fig. 6). The water film can be removed by vacuum sipping. Zeolite

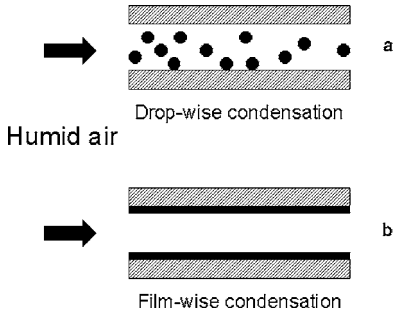


Figure 6. Schematic illustration of the use of a hydrophilic zeolite coating on a heat exchanger surface to change condensation mode from drop-wise without zeolite film (a) to film-wise with zeolite film (b).

coatings were shown to be highly hydrophilic with advancing contact angle of 0–15 degrees and a receding contact angle of 0 degrees. The adhesion was 5 B under ASTM method 3395.

In-situ crystallization was used and clearly preferred as it can coat surfaces of complex shape and in confined spaces. In addition, it was a low temperature process (e.g., 165 °C) and compatible with most metal alloys. Another significant advantage for a zeolite coating is that an antimicrobial characteristic can be added readily through ion exchange of metal ions such as silver. Antimicrobial function is desirable for hydrophilic coatings as damp surfaces tend to grow bacteria and fungi. Furthermore, once depleted, the antimicrobial capability of zeolite coatings can be easily regenerated by another ion-exchange without the need of a potentially costly recoating.

4.9. Heat Pump

It is well known that water adsorption in zeolites releases heat while its desorption absorbs heat. This phenomenon has been exploited for the design of a heat pump for cooling applications using waste heat or solar energy. At present, zeolite pellets are used and the heat and mass transfer are inefficient, leading to bulky pumps. The utilization of zeolite coatings could eliminate these limitations. Zeolite coatings on metal substrates have been claimed for enhancing transport of sensible and latent heat from warm air to cool air. Moisture adsorption/desorption was exploited. The coatings were required to be capable of adsorbing/releasing moisture from/to an air stream, but not capable of adsorbing contaminants from the air stream [144]. There have been renewed interests in zeolite coating heat pumps recently [145, 146]. A new *in-situ* crystallization method for zeolite coatings was developed which suppresses the reaction in the bulk and promotes it on the substrate by applying a temperature difference between the reaction mixture and the substrate. The substrate was heated directly, while the reaction mixture was kept at a lower temperature by means of a water bath. The method was tested for zeolite 4A coatings on stainless steel plates from a clear aluminosilicate synthesis solution. Impurity seemed to be a concern, and a final treatment that involved an increase in the temperature of the water bath could be applied to the coatings at the end of their preparation in order to remove the impurities [145].

In another study by the same group, the thermal and mechanical stabilities of the coatings were determined by the application of consecutive heating and cooling cycles and ultrasonic treatments. Although ultrasonic treatments resulted in relatively higher mass losses, the stabilities of the zeolite coatings synthesized on stainless steel plates were observed to be only moderately affected by the very rapid and consecutive heating and cooling cycles applied. No difficulty was observed for desorption of water vapor out of the zeolite layer [146].

4.10. Thermoelectrics

Thermoelectrics are all solid-state devices that can directly convert heat to electricity (power generation) or use electricity to pump heat from cold to hot (refrigeration), both without using any moving parts and hazardous compressed fluids. Thermoelectric devices are lightweight, compact, and with extremely fast response, and therefore are particularly useful for applications such as microelectronics and semiconductor lasers where conventional compression-based cooling is very difficult or impossible. At present, applications of thermoelectric devices are limited because of their low efficiency. The efficiency of a thermoelectric device is usually characterized by a dimensionless figure of merit, ZT. Current thermoelectric materials have a $ZT = 1$. A thermoelectric cooler with a $ZT = 1$ operates at only 10% of Carnot efficiency, whereas a thermoelectric cooler with a $ZT = 4$ would achieve 30% of Carnot efficiency, which is comparable to the efficiency of a home refrigerator.

Very recently, theoretical calculation for isolated single nanowires predicts over a 10-fold increase in the figure-of-merit. In order to realize these size effects, wires must be synthesized with diameters less than the thermal de Broglie wavelength (typically less than 10 nm), which is smaller than is currently accessible by lithography techniques. Also, in order to transfer meaningful amounts of thermal energy, a practical device must consist of an array of a large number of nanowires in parallel. One way to potentially synthesize these nanowire arrays is to use the pores of microporous and mesoporous materials as a mold to template the diameter and orientation of wires of suitable thermoelectric materials. A recent review has been published discussing the relevant physics and derived a model to assess the feasibility of using microporous and mesoporous frameworks in the fabrication of thermoelectric devices [147]. The model accounted for the possible deleterious effects of thermal conduction through the microporous or mesoporous framework and the presence of bulk thermoelectric material (that may result from defects in the microporous or mesoporous framework) in parallel with the wires on the thermoelectric figure-of-merit. Simulation results were reported for SBA-15, MCM-41, and zeolites VFI, LTL, and LTA embedded with Bi_2Te_3 nanowires. The results showed that the microporous zeolite frameworks may yield figures-of-merit much larger than one, while thermal conduction through most mesoporous frameworks reduces the figure-of-merit below the level currently available using traditional thermoelectric materials.

5. COLLECTIONS OF ZEOLITE FILMS

A large number of articles have been published on zeolite films. This section was intended to be a comprehensive collection of these reports up to 2002 although accidental omission is inevitable. To avoid duplication, papers that were used as examples to illustrate the synthesis strategies and applications in previous sections are not included in this section except in some special cases. This section begins with three most widely studied zeolite films: MFI, FAU, and LTA. The rest is in alphabetical order according to their three-letter codes. This format makes it easy for future expansion of this review as the field of zeolite films evolves.

5.1. MFI (ZSM-5 and Silicalite)

5.1.1. Separation Membranes

Membrane Preparations. MFI membranes have been prepared by *in-situ* crystallization on a variety of supports including porous alumina and sol–gel derived mesoporous yttria doped zirconia [148], a TiO_2-coated porous stainless steel support [149], porous sintered glass tube [150], porous alumina and other substrates [151, 152], porous stainless steel and alumina tubes using an alkali-free synthesis mixture [153]. The *in-situ* crystallization method could be assisted by an electrophoretic technique to attract zeolite/gel particles to the substrate surface [154]. The surface properties of the porous stainless steel support is very important and the effects of surface functional groups (-SiO_2, -OH, and -COOH) were simulated using adsorbed surfactants [155]. In an effort to have a separation module without dealing with the sealing problem, MFI membranes were grown within stainless steel membrane modules and the MFI layers remained thermomechanically stable upon calcination at 400 °C in air [156].

MFI membranes have also been prepared by seeded growth methods on a variety of supports, including inside the pores of porous alumina support tubes [157], porous alumina tubes with seeding under ultrasonication [158, 159], porous alumina disks and nonporous substrates with preferential orientation [160]. MFI membranes prepared by seeded growth on mesoporous alumina support were compared with those by *in-situ* crystallization for their performance in pervaporation and steam permeation of water/iso-propanol and water/methanol mixtures [161]. The seeded growth process of polycrystalline MFI films was simulated in two spatial dimensions using a front tracking technique [162].

An interesting combination of seeded growth and a vapor phase transport method was used to prepare silicalite membranes. A silica layer was first deposited on the porous support and followed by a silicalite nanocrystal layer. Then the coated substrate was treated with steam in a closed reactor [163, 164].

Membrane Separations. MFI membranes have been demonstrated for several classes of separations based on adsorption, diffusion, molecular sieving, or a combination of these. The first major class is separation of organic/water mixtures through pervaporation. Selectivity

for pervaporation is mainly determined by the hydrophobicity/hydrophilicity of the membrane. Demonstrated pervaporations include separation of mixtures of methanol/MTBE [165], 1,3-propanediol/glycerol and glucose/water [166], ethanol/water [167, 168], methyl ethyl ketone/water, methanol, ethanol, and acetone/water [169].

The second major class is separation of hydrocarbons such as xylene isomers and linear/branched hydrocarbons. Examples include pervaporation of pure xylene isomers and their binary mixture through alumina-supported MFI membranes [170], separation of C_4, C_6, and xylene isomers through a MFI membrane prepared by seeded growth [171], permeation of n-octane, isooctane, and n-hexane vapors through silicalite membranes on porous alumina [172], n-hexane/2,2-dimethylbutane [173, 174]. One interesting development is the use of an MFI composite tubular membrane to improve the stability of liquid membranes [175].

The third one is separation of fixed gases and hydrogen. Examples here include separation of CO_2 and CH_4 [176], binary mixtures of CH_4-N_2, CO_2-N_2, and CO_2-CH_4 [177], a CO/air mixture at very low CO concentration (160 ppbv) [178], hydrogen in a Town gas mixture [179], and hydrogen in a refinery gas mixture [180].

Many separation experiments were carried out aiming to understand the permeation behavior. Studied were effects of operating conditions on the flux such as temperature, feed pressure, sweep gas flow rate, and orientation of the membrane [181, 182], pressure drop and sweep gas versus vacuum on the permeate side [183], influence of membrane microstructure [184], chemical nature of the sweep gas and orientation of the composite membrane [185–187].

Membrane Transport Theory. Transport in zeolite membrane is a rather complex process involving adsorption and diffusion. The presence of defects adds extra complexity to the problem. Significant amount of theoretical work has been carried out aiming to understand and perhaps eventually predict the permeation of molecules. Maxwell–Stefan formulation appears to be widely used either alone or together with other techniques [188–190]. Diffusion of methane, n-butane, n-hexane, isobutane, and 3-methylpentane was studied, and Maxwell–Stefan equations were used for the interpretation of the results [191]. Models based on Maxwell–Stefan formulation were also used to describe transport of supercritical fluids in zeolite materials [189].

A combination of adsorbed solution theory to describe the mixture sorption, and Maxwell–Stefan equations for diffusion have been used for modeling transport through zeolite membranes [192]. Similarly, a combination of configurational bias Monte Carlo simulations for pure component and mixture isotherms and Maxwell–Stefan theory for diffusion provided a nearly quantitative explanation of the experimental data on separation of n-hexane from 2,2-dimethylbutane (DMB) [193].

Other techniques included simple models based on surface diffusion, activated gaseous diffusion and viscous flow [194, 195], molecular simulation [196], dual control volume grand canonical molecular dynamics [197], and dual ensemble Monte Carlo simulation for pervaporation of a Lennard–Jones (LJ) fluid mixture [198, 199]. Models addressing nanoscopic defects were also developed [200].

5.1.2. Nonseparation Films

MFI films were formed by *in-situ* crystallization on a variety of substrates such as cordierite modules [201], activated carbon with suitable nucleation sites introduced by adding hydrophilic montmorillonite clay [202]. The change in morphology of ZSM-5 films during steaming at 443 K was studied [203]. Large differences were observed in morphology and an Si/Al ratio between outer and inner sides of the films on a nonporous substrate [204]. The effects of aging of source materials were studied for films on quartz surfaces [205]. A model for nucleation and growth of silicalite crystals was proposed [206]. Free-standing films were synthesized under elevated gravity (16 G) [207].

Seeded growth methods were also widely used for forming MFI films on polymer and glasses [208], carbon, ZrO_2, Al_2O_3, mullite, ceramic and pyrex glass fibers [209], gold [210, 211]. The out-of-plane grain orientation could be manipulated [212]. Micrometer-sized zeolite particles were found to form relatively stable particle monolayers at the air-water interface when spread from a chloroform dispersion. The transfer of the floating thin film onto a solid substrate via the Langmuir–Blodgett (LB) technique proved to be possible [213].

5.2. FAU (X, Y, Faujasite)

Zeolite X films were formed by seeded growth method on TiN-coated silicon wafers and porous stainless steel disks using laser ablation for seed layer deposition [214], and on porous alumina tubes showing permselectivity of CO_2 to N_2 of 20–100 [45]. Na-X membranes by *in-situ* crystallization contain adsorbed azobenzene (AZB) were shown to have photoswitchable permeation properties due to the trans-cis photoisomerization of AZB [215].

Zeolite Y membranes were synthesized by *in-situ* crystallization on the outer surface of a porous alumina tube and ion-exchanged with KCl, RbCl, or CsCl. The CO_2/N_2 selectivity for the equimolar mixture was in the order of Rb = K > Cs > Na [216]. In another study by the same group, zeolite Y membranes were ion-exchanged using Pt, Ru, Rh, Co, Ni, Cu, or Ag. The metal-loaded Y membranes were shown to have activity for CO oxidation with PtY membrane being the most active [217]. The effects of cations of alkaline and alkaline earth elements were studied and K-Y membranes gave the highest CO_2/N_2 selectivity [218].

Permeation through Y membranes was calculated by density function theory. The results successfully predict the experimental observation of selective permeation of gaseous molecules in the order $C_2H_6 < CH_4 < SF_6 < CO_2 < N_2$ [219]. The separation of carbon dioxide from the carbon dioxide and nitrogen mixture was calculated by molecular dynamics (MD) method [220]. Zeolite Y coatings on stainless steel were achieved. The bonding with the support and thus the coverage are improved by an oxide layer on the surface formed upon calcination at high temperature [221].

FAU membranes with an Si/Al ratio of 1.3–1.8 were prepared by seeded growth on porous alumina disks. The selectivity for N_2/CO_2 and CH_4/CO_2 were 8.4 and 3.5, respectively [222]. FAU membranes with Si/Al about 1–1.5 were synthesized on porous alumina by *in-situ* crystallization and seeded growth, and the mixtures examined (and

the corresponding equimolar mixture separation factors) were benzene/cyclohexane (160), benzene/n-hexane (144), toluene/n-heptane (45), propylene/propane (6.2), and ethylene/methane (8.4) [223]. Oriented FAU films were achieved on compositionally different, micrometer-sized EMT zeolite crystals [224, 225].

5.3. LTA (A)

The preparation of zeolite A films began with a report on free-standing film [226]. Thereafter, zeolite A membranes were commonly prepared by seeded growth method and used for pervaporation of water/organic mixtures [227–230]. It was shown that surface seeding could accelerate the formation of NaA on the support and inhibit the transformation of NaA into other zeolites. In addition good selectivities for H_2/n-C_4H_{10} and O_2/N_2 were achieved [231, 232]. Preferred orientation was demonstrated [233–236].

Zeolite A membranes have also been prepared by *in-situ* crystallization using a local heating method [145] intended for heat pump applications. Several other interesting applications have been demonstrated for A membranes including sensors [119, 237], and reverse osmosis [238]. A molecular simulation scheme was also developed for osmosis and reverse osmosis [239].

5.4. AFI (SAPO-5 and ALPO-5)

5.4.1. AFI (SAPO-5)

SAPO-5 films and membranes were prepared on a variety of substrates such as steel and Cu [240], Teflon slab [241], filter paper [242], and anodic alumina [243]. Separation of aromatics has been studied [244] and preferred orientation has been achieved [243].

5.4.2. AFI (ALPO-5)

There are many studies on $ALPO_4$-5 films and membranes. Some of these are pure, supported zeolite membranes and others are polymer-zeolite composite membranes [245, 246] and zeolite imbedded in nonporous metals [247]. For supported zeolite films and membranes, both *in-situ* crystallization [248] and seeded growth using laser ablation for seed deposition [46, 249] were used. Several applications were demonstrated including separation [245, 246], and sensors [250, 251]. There are also reports on oriented films on Au [251] and oriented membranes on anodized alumina substrate [252–254], oriented crystals embedded into a metallic grid and fixed by galvanic Ni deposition [255, 256].

$ALPO_4$-5 membrane reactors have also been studied and the membranes were evaluated as O_2 distributors in a passive membrane reactor for the partial oxidation of alkanes [257]. $ALPO_4$-5 films could also have applications in nonlinear optics [258]. A simplified lattice gas model was used to describe adsorbate passage inside $ALPO_4$-5 pores [259], and an atomistic model was developed for single gas permeation in $ALPO_4$-5 membranes [260].

5.5. ANA (Analcime)

Polycrystalline ANA films were prepared by *in-situ* crystallization. The films were about 200 μm thick and were deposited on a porous alumina support at temperatures above 160 °C. The films were thermally stable at temperatures up to 700 °C, and no structural changes were observed after treatment in nitric acid solution below 0.01 M. An ion-sieve capability was demonstrated [261], and treatment of radioactive waste using these films was suggested [262].

ANA films were also prepared by vapor phase transport method on porous alumina substrate. Dissolution of aluminum was suggested to be responsible for the formation of high aluminum zeolites like ANA. When the support was coated with silica, high-silica zeolites such as FER and MFI were formed [263].

5.6. BEA (Beta)

Zeolite BEA membranes were prepared by *in-situ* crystallization on porous stainless steel [264] and ceramic filter [265]. Separations of gases (H_2, N_2, n-, and i-butane) and liquids THF/H_2O, iPrOH/H_2O, tBuOH/H_2O, 1,3-propanediol/H_2O, glycerol/H_2O, and 1,3-propanediol/glycerol/H_2O [264], and PhOH/H_2O and C_6H_6/H_2O [265] were studied.

Zeolite BEA thin films were grown on a Ta-substrate using colloidal crystals as seeds [266]. Zeolite BEA coatings on structured packings were used for catalytic distillation to prepare octane booster t-BuOEt (ETBE) [267] and to improve the etherification of methanol and isobutene [268]. Zeolite BEA coatings were also prepared on polystyrene beads and then the polymer beads were removed to generate zeolite hollow spheres [269].

5.7. CHA (Chabazite and SAPO-34)

5.7.1. CHA (Chabazite)

Free-standing chabasite films were prepared on walls of polypropylene reaction bottles. Relatively large pieces can be readily peeled off from the walls of the polypropylene bottle [270]. Supported pure-phase chabasite membranes have not been reported. A mixed-phase zeolite membrane (mordenite/ZSM5/chabazite) was prepared on alumina tubular supports by *in-situ* crystallization. This membrane was used for study of permeation of single gases (N_2 and n-butane), ternary mixtures containing water, alcohol (methanol, ethanol, or propanol), and a permanent gas (O_2) [271, 272], and a binary mixture of n-butane/H_2 and n/i-butane [273]. Fuel cell membranes containing chabasite were also proposed and studied [274]. Heteroepitaxial growth of chabazite on sodalite was also achieved [275].

5.7.2. CHA (SAPO-34)

SAPO-34 membranes have been primarily prepared by *in-situ* crystallization [276] on alumina supports [277] and ceramic monolith [278]. Several separations were attempted and these included single gases and mixtures of H_2/CH_4, CO_2/CH_4, H_2/N_2, and CO_2/N_2 [277, 279]. The effects of humidity on these membranes were studied [280].

5.8. DON (UTD-1)

Highly oriented films of high-silica DON were prepared using the seeded growth method. The seed layer was deposited by pulsed laser ablation of as-synthesized DON

crystals onto polished Si wafers. Densely packed aggregates of elongated DON crystals grew from the ablated surface with the zeolite channels oriented perpendicular to the substrate [281]. This technique was later extended to small three-dimensional objects. Specifically, oriented DON zeolite films were prepared on the surface of glass and steel beads by first vibrating the spheres in a laser-generated plume [282].

5.9. ERI/OFF (T)

Zeolite T membranes were used for pervaporation-aided esterification of acetic acid with ethanol to prepare ethyl acetate. Almost complete conversion was reached within 8 h when initial molar ratios of alcohol to acetic acid were 1.5 and 2 [283, 284].

5.10. FER (Ferrierite)

FER membranes were prepared on porous alumina by vapor-phase transport method using ethylenediamine (EDA), triethylamine (TEA), and H_2O as vapor sources [285]. The formation mechanism was studied and pervaporation of bezene/xylene, and m-xylene/1,3,5-triisopropylbenzene mixtures was performed [286–288].

5.11. KFI (P)

Zeolite KFI membranes were synthesized by *in-situ* crystallization on porous alumina supports. Permselectivity as high as 102 was obtained for H_2/SF_6. At room temperature, the permselectivities for H_2/Ar and CH_4/Ar were 5.29 and 2.36, respectively. The apparent activation energies for permeation of H_2, CH_4, and Ar in 25–65 °C were estimated to be 11.1, 9.1, and 15.2 kJ/mol, respectively. In a dry atmosphere, the crystal structure gradually transformed from a cubic to a tetragonal phase at temperatures higher than 65 °C. The transformation did not result in drastic changes in gas permeance and permselectivity, indicating that KFI films were still continuous. KFI membranes were more stable in a humid atmosphere [289].

5.12. MOR (Mordenite)

in-situ crystallization was used to prepare MOR membranes on porous alumina and stainless steel [290] for separations of several mixtures such as water from water/He, water/HC, water/HC/He mixtures (HC for methane, propane, or n-butane) [291], and water/iso-propyl alcohol [292]. In another study, a mixed-phase zeolite membrane containing MOR/MFI/CHA was prepared on alumina tubes and used to separate ternary mixtures containing water, alcohol (methanol, ethanol, or propanol), and O_2 [271]. The membrane formation mechanism has been studied [293] and separation of aromatics (p-xylene, o-xylene, and benzene) was demonstrated [244]. Free-standing membranes were prepared [294].

The seeded growth method was used to prepare MOR membranes on porous alumina supports for pervaporation of a water/iso-propanol mixture [295]. The vapor phase transport method was demonstrated on porous alumina for separations of benzene over p-xylene [263]. Parallel diffusion models were proposed for permeation of H_2, He, CH_4, N_2, O_2, and CO_2 [296]. Preferred orientation was also achieved [297].

Pulsed laser deposition was used to deposit MOR films for gas sensor applications. These films were comprised of small crystallites embedded in an amorphous matrix. The amorphous material was microporous having pore sizes on the order of those in the parent zeolites. These pores were oriented perpendicular to the substrate. Quartz crystal microbalances coated with the material have been tested for gas uptake and were shown to discriminate between isomers of small molecules [107, 298].

MOR layers were grown on a metallic Cr surface by vapor phase transport for catalytic applications [299]. Honeycomb supports were coated with a mixture of zeolites ion-exchanged with rare earth elements and zeolites ion-exchanged with Cu or Co for improved catalytic NOx removal. The use of MOR was suggested [300]. Structured packings coated with BEA and MOR were used to prepare octane booster t-BuOEt (ETBE) in a liquid phase batch system [267].

Electrochemical applications of MOR films were demonstrated. For example, Cs-selective electrodes were prepared from MOR crystals embedded in an epoxy resin [301]. MOR composite films were also suggested as proton conductors for proton exchange membrane (PEM) fuel cell applications [302, 303].

5.13. MTW (ZSM-12)

High-silica MTW coatings were synthesized by *in-situ* crystallization using a new two-silica method [142] on Al-alloys (Al-2024-T3 and Al-6061-T4) and stainless steel (SS-304). All coatings were continuous and had excellent adhesion. The as-synthesized zeolite coatings were tested by a dc polarization technique and found to be highly corrosion-resistant in severely corrosive acidic (0.5 M H_2SO_4) and alkaline (0.1 M NaOH) media [141].

5.14. MWW (MCM-22)

MWW films were grown on glass substrates by *in-situ* crystallization [304, 305]. The thin disk-like MWW crystals were found to be perpendicular to the substrate surface in the early stage of crystallization and then grew thicker to form a continuous film. It was suggested that the fast growth rate in the layer direction was responsible for the observed orientation.

MWW films were prepared on TiN-coated silicon wafers, polished graphite, and porous stainless steel frits by seeded growth. The seed layer was introduced by pulsed laser ablation of as-synthesized layered MWW. The films were partially oriented with a-direction largely normal to the substrate surface. The oriented MWW membrane grown on porous stainless steel was evaluated for room temperature separation of an alcohol/water mixture and a selective permeation of water was observed [306].

5.15. ETS-4

Since Engelhard titanium silicates have octahedral coordination in addition to tetrahedral coordination, they are officially non zeolite-type materials. Nonetheless, ETS films could have interesting membrane and film properties, and therefore are included in this review. A combination of *in-situ* crystallization and a seeded growth method was used to prepare ETS-4 supported films and membranes. It involved formation of a seed layer on a support by *in-situ* crystallization followed by seeded growth. The synthesized films and membranes exhibited b-out-of-plane preferred orientation that corresponded to the eight-member ring channels of ETS-4 oriented perpendicular to the substrate. Membranes consisting of thin intergrown films of Na-ETS-4 on porous TiO_2 supports were highly H_2O permselective for 1:1 H_2O/ethanol mixtures. The selectivity of the ETS-4 membranes was similar to the highest reported for Na-X and Na-Y membranes, and approximately 10 times lower than the selectivities of zeolite Na-A prepared in the same lab and up to 100 times lower than the highest selectivities reported in the literature for zeolite Na-A membranes. The flux of H_2O through Na-ETS-4 membranes was comparable or higher than that through zeolite A, X, and Y membranes [307].

6. RECENT DEVELOPMENTS

6.1. Oriented Zeolite Films

Almost all applications for zeolite films could benefit from a preferred orientation. This is especially critical for zeolites with anisotropic pore systems. Noncontinuous zeolite layers with preferred orientation have been prepared [34, 57, 254, 308, 309], but this section focuses on continuous oriented polycrystalline films. Examples include b-oriented MFI by *in-situ* crystallization on MFI [41, 44], c-oriented MFI [28] and LTA [233] by seeded growth [310], and MFI [70] and MOR [295] by vapor phase transport.

One particularly interesting b-oriented MFI film is shown in Figure 7. It is a continuous monolayer film (0.4–0.5 μm thick) by *in-situ* crystallization on metal substrates (stainless steel and Al alloy) [41, 44] or on silicon wafer [138]. Since MFI has straight channels in the b-direction, this film is well suited for membrane application. It has been found that there is positive thermal expansion in a- and c-direction and negative expansion in b-direction, thus a b-oriented MFI film helps to relieve thermal stress. The thinness of the film is also important for relieving thermal stress. Also because this is a monolayer film, it does not have any grain boundaries in the direction perpendicular to the substrate, leading to minimized resistance for mass transport. This film could also be useful for sensor, magnetic, optical, thermoelectric applications as the template for nanowire arrays.

6.2. Zeolite Nanocrystals as Building Blocks

Zeolite nanocrystals have been demonstrated to be versatile building blocks for constructing hierarchical porous structures [311–317]. The use of nanoparticles as building blocks allows mild processing conditions and innovative strategies to be used that would have been difficult or impossible otherwise with the conventional hydrothermal

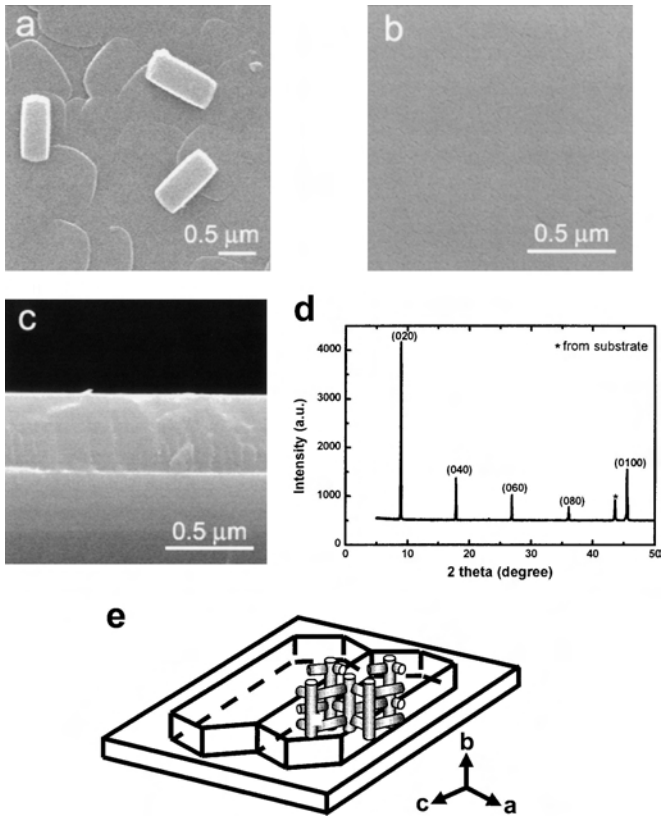


Figure 7. (a, b, c) SEM micrographs of silicalite films on silicon wafer by *in-situ* crystallization, (a) before polishing, top view, (b) after polishing, top view, (c) after polishing, cross-sectional view; (d) XRD pattern of the silicalite film showing that all the straight channels are vertical to the silicon substrate; (e) a schematic showing the orientation of the crystals versus the substrate (after [41, 44, 138]).

synthesis methods such as *in-situ* crystallization, seeded growth, and vapor phase transport. The fabrication of three-dimensional structures with designed shapes via gel-casting [315] and transparent zeolite films [312] (Fig. 8) are examples that clearly illustrate the utilities and versatility of this approach.

Specifically for the preparation of zeolite films, zeolite nanoparticles could be used directly to form zeolite films by self-assembly [311]. They could also be combined nicely

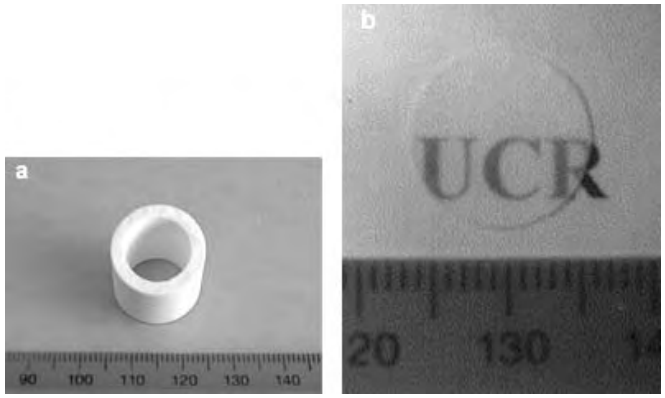


Figure 8. Zeolite structures constructed by using zeolite nanocrystals as building blocks (after [312, 315]).

with seeded growth and the vapor phase transport method to produce high-quality polycrystalline zeolite films [28, 163]. There is clear evidence that small zeolite nanoparticles are preferred for producing compact continuous films.

Another important application for zeolite nanaoparticles is the preparation of zeolite matrix composite films. Micrometer-sized crystals have been used in these types of film using techniques such as wash-coating and sol–gel processing. However, the use of nanoparticles in these films appears to offer unparalleled advantages in terms of achievable film thickness and uniformity and processability. The low-k dielectric films [30, 139] and polymer-zeolite nanocomposite membranes [33, 318] are two examples. The efforts for polymer-zeolite nanocomposite films were primarily driven by gas separation applications. Zeolite membranes have high selectivity and high thermal and chemical stability. But it is difficult to develop a high density module such as hollow fibers that are commonly used by polymer membranes. On the other hand, polymer membranes can be easily processed but have relatively low selectivity. Therefore, it has been an outstanding goal to combine the excellent properties of these two materials.

Another interesting application is microreactors. Zeolite adsorbents and catalysts have been an important part for the conventional reaction-separation system. As the efforts to miniaturize reactors evolve, using zeolite nanocrystals as building blocks could be an ideal way to introduce zeolites into these microreactor systems.

6.3. Patterned Zeolite Films

The ability to pattern zeolite films is critical for many of their potential applications, especially optical, magnetic, and electronic applications that have not been adequately explored. The first patterned zeolite film was achieved by using zeolite nanocrystals as building blocks combined with soft lithography [311] (Fig. 9). Convection-assisted assembly was also demonstrated for a generation of surface-patterned zeolite films although the patterns were not very regular [31] (Figure 9). Recently, using organic linkers, Yoon et al. have demonstrated that microcrystals can also be patterned and oriented. Although these films have very sophisticated patterns, they are not continuous and adhesion is not very strong. There is a critical need to fabricate oriented continuous films on a variety of substrates.

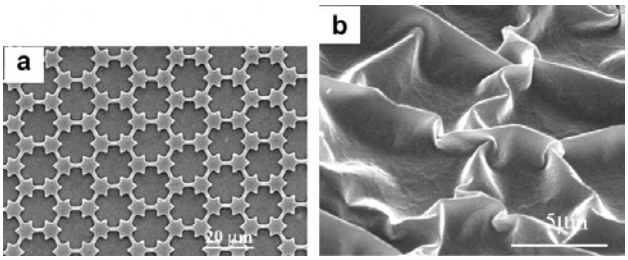


Figure 9. Patterned zeolite films by using (a) soft-lithography and (b) convection-assisted assembly of zeolite nanoparticles (after [31, 311]).

7. CONCLUDING REMARKS

The field of nanostructured zeolite films has come a long way during the last decade from almost nonexistence in the early 1990's to a vibrant field today with many applications demonstrated to be promising, and one commercial application already realized. Although scattered activities can be traced back to the 1960's, the key to this success has been the persistent work of a large number of researchers around the world during the last 10 years. The single biggest driver behind this collective effort so far appears to be the potential application of zeolite films as membranes and membrane reactors. But nonconventional applications (i.e., nonseparation and noncatalysis) have been quickly added and these new applications promise to offer even more exciting opportunities. Table 1 shows some of the properties of nanostructured zeolite films. As shown in this review, zeolite based-nanostructured films could take on several different configurations and be prepared by a number of synthesis methods. Current strategies for oriented films are highly specific for a particular zeolite or a particular orientation. More general methods for preparing oriented films are still critically needed. The use of nanoparticles as building blocks has broadened our capability significantly and will continue to bring exciting opportunities to the field. Although not covered in this review, additional new applications such as optical, magnetic and electric devices are being and will continue to be explored. For these device applications, the ability of patterning by simple techniques will be crucial.

Table 1. Physical properties of nanostructured zeolite films.

Nanostructured zeolite films	Dielectric constant	Elastic modulus (GPa)	Corrosion current (A/cm^2)	Contact angle (Degree)	Micropore size (nm)	Mesopore size (nm)	Ref.
MFI (silicalite) by *in-situ* crystallization	2.7	30–40	10^{-8}	—	0.55	N/A	[138]
MFI (silicalite) from nanocrystals	1.8	—	—	—	0.55	17	[138, 311]
MFI (silicalite) from nanoparticles	2.1	16–18	—	—	0.55	2.6–2.8	[30, 31]
MFI (ZSM-5) by *in-situ* crystallization	—	—	10^{-8}–10^{-12}	0–15 advancing 0 receding	—	N/A	[42]
BEA (pure-silica) by *in-situ* crystallization	—	—	10^{-8}	—	—	N/A	[141]

GLOSSARY

Calcination A process where a sample is heated in an oxidizing environment.

Chemical mechanical polishing (CMP) A semiconductor processing step that is used to level the interconnect layers.

Chemical vapor deposition (CVD) A commonly used semiconductor process for deposition of thin films.

Low dielectric constant material (Low-k material) Dielectric material is used in computer chips to insulate the wirings among the transistors.

Pressure swing adsorption (PSA) This is a process that separates gases by selective adsorption of a component in a mixture to an adsorbent such as zeolites.

Structure directing agent (SDA) These are organic molecules that are used to direct the synthesis to a desired structure.

Tetrahedrally coordinated atom (T-atom) In zeolites T-atoms refer to atoms such silicon and aluminum that are coordinated by four oxygen atoms in the crystal structure.

Tetraethylorthosilicate (TEOS) This is a commonly used silica source for zeolite synthesis.

Thermoelectric Thermoelectric materials can provide solid state refrigeration using electricity or generate electricity from a temperature difference.

Zeolite A class of porous crystals with uniform molecular-sized pores. Zeolites can be found in nature and can also be synthesized in laboratory.

ACKNOWLEDGMENTS

The authors wish to thank their sponsors for providing financial support and their colleagues in the Yan Group at UC-Riverside for contributing to their work which is included in this article.

The sponsors are as follows: AlliedSignal/Honeywell International, Advanced Micro Devices, Asahi Kasei Corporation, Pacific Fuel Cells Corporation, The Technology for Sustainable Environment (TSE) program of the U.S. Environmental Protection Agency, The Strategic Environmental Research and Development Program (SERDP) of the U.S. Department of Defense, University of California—Semiconductor Manufacturing Alliance Research and Teaching (UC-SMART), University of California—Toxic Substance Research & Teaching Program (UC-TSR&TP), University of California—Energy Institute (UC-EI), California Energy Commission, and Riverside Public Utilities.

The authors' colleagues include: Dr. Z. Wang, Dr. X. Cheng, Dr. A. Mitra, Dr. L. Huang, Dr. S. Li, Dr. C. Wang, Dr. H. Luo, T. Cao, D. Beving, B. Holmberg, Z. Li, A. Avila, J. Zhao, M. Waje, and C. Demmelmaier.

REFERENCES

1. D. W. Breck, "Zeolite Molecular Sieves: Structure, Chemistry, and Use." Wiley, New York, 1974.
2. W. M. M. Ch. Baerlocher and D. H. Olson, "Atlas of Zeolite Framework Types," 5th ed. Elsevier, Amsterdam, 2001.
3. R. Szostak, "Handbook of Molecular Sieves." Van Nostrand Reinhold, New York, 1992.
4. J. Weitkamp, U. Weiss, and S. Ernst, *Stud. Surf. Sci. Catal.* 94, 363 (1995).
5. C. Song, *Stud. Surf. Sci. Catal.* 113, 163 (1998).
6. T. Bein, *Cur. Opin. Solid S. M.* 4, 85 (1999).
7. A. J. W. Angell, A. Dorset, and D. W. York, *Eur. Pat. Appl.* 627484 (1994).
8. J. A. Kosal and E. P. Plueddemann, U.S. Patent 4534880, 1985.
9. P. Gorlin, L. Zyzyck, and G. Jakubicki, U.S. Patent 6191089, 2001.
10. P. A. Gorlin and S. Phillips, U.S. Patent 6025315, 2000.
11. P. W. Appel, T. J. C. Arts, C. A. Lanceley, and T. J. Osinga, *PCT Int. Appl.* 9734980 (1997).
12. P. W. Appel, T. J. C. Arts, J. V. Boskamp, A. P. Chapple, F. C. P. M. Dobbe, C. M. B. Joyeux, C. A. Lanceley, E. L. M. Lempers, and T. J. Osinga, *PCT Int. Appl.* 9734979 (1997).
13. M. P. Astley and W. D. Emery, *PCT Int. Appl.* 0181528 (2001).
14. B. Lenoble and G. Demaret, *Eur. Pat. Appl.* 997180 (2000).
15. J. C. Jansen, J. H. Koegler, H. Van Bekkum, H. P. A. Calis, C. M. Van Den Bleek, F. Kapteijn, J. A. Moulijn, E. R. Geus, and N. Van Der Puil, *Micropor. Mesoporous Mat.* 21, 213 (1998).
16. J. C. Jansen, J. M. van dar Graff, N. van der Puil, S. B. G. Seijger, and S. P. J. Smith, "Proceedings of the 12th International Zeolite Conference," 1999, 1, 603.
17. K. C. Jansen and E. N. Coker, *Cur. Opin. in Solid State Mater. Sci.* 1, 65 (1996).
18. T. Bein, *Chem. Mater.* 8, 1636 (1996).
19. K. K. Sirkar, *Chemical Engineering Communications* 157, 145 (1997).
20. A. Tavolaro and E. Drioli, *Adv. Mater.* 11, 975 (1999).
21. M. Tsapatsis, G. Xomeritakis, H. Hillhouse, S. Nair, V. Nikolakis, G. Bonilla, and Z. Lai, *Cattech* 3, 148 (2000).
22. L. Cot, A. Ayral, J. Durand, C. Guizard, N. Hovnanian, A. Julbe, and A. Larbot, *Solid State Sciences* 2, 313 (2000).
23. J. Caro, M. Noack, P. Kolsch, and R. Schafer, *Microporous Mesoporous Mater.* 38, 3 (2000).
24. M. E. Davis, *Nature* 417, 813 (2002).
25. F. Schuth and W. Schmidt, *Adv. Mater.* 14, 629 (2002).
26. Y. Yan, M. E. Davis, and G. R. Gavalas, *Ind. Eng. Chem. Res.* 34, 1652 (1995).
27. W. O. Haag and J. G. Tsikoyiannis, U.S. Patent 5019263, 1991.
28. M. C. Lovallo and M. Tsapatsis, *AIChE J.* 42, 3020 (1996).
29. M. Matsukata, N. Nishiyama, and K. Ueyama, *Chem. Commun.* 339 (1994).
30. Z. Wang, A. Mitra, H. Wang, L. Huang, and Y. Yan, *Adv. Mater.* 13, 1463 (2001).
31. H. Wang, Z. Wang, L. Huang, A. Mitra, and Y. Yan, *Langmuir* 17, 2572 (2001).
32. H. Wang, L. Huang, B. A. Holmberg, and Y. Yan, *Chem. Commun.* 1708 (2002).
33. H. Wang, A. Holmberg Brett, and Y. Yan, *J. Mater. Chem.* 12, 3640 (2002).
34. S. Feng and T. Bein, *Nature* 368, 834–6 (1994).
35. S. Feng and T. Bein, *Science* 265, 1839 (1994).
36. S. Y. Choi, Y.-J. Lee, Y. S. Park, K. Ha, and K. B. Yoon, *J. Am. Chem. Soc.* 122, 5201 (2000).
37. K. Ha, Y.-J. Lee, D.-Y. Jung, J. H. Lee, and K. B. Yoon, *Adv. Mater.* 12, 1614 (2000).
38. K. Ha, Y.-J. Lee, H. J. Lee, and K. B. Yoon, *Adv. Mater.* 12, 1114 (2000).
39. A. Kulak, Y. S. Park, Y.-J. Lee, Y. S. Chun, K. Ha, and K. B. Yoon, *J. Am. Chem. Soc.* 122, 9308 (2000).
40. Y. Yan, M. Tsapatsis, G. R. Gavalas, and M. E. Davis, *Chem. Commun.* 227 (1995).
41. Z. Wang and Y. Yan, *Chem. Mater.* 13, 1101 (2001).
42. X. Cheng, Z. Wang, and Y. Yan, *Electrochem. Solid-State Lett.* 4, B23 (2001).

43. Y. Yan, U.S. Patent 6,500,490, 2002.
44. Z. Wang and Y. Yan, *Microporous Mesoporous Mater.* 48, 229–238 (2001).
45. K. Kusakabe, T. Kuroda, A. Murata, and S. Morooka, *Ind. & Eng. Chem. Res.* 36, 649 (1997).
46. K. J. Balkus, Jr., L. J. Sottile, S. J. Riley, and B. E. Gnade, *Thin Solid Films* 260, 4 (1995).
47. Z. Li, S. Li, and Y. Yan, Unpublished results, 2002.
48. M. D. Patil and I. M. Lachman, *ACS Symp. Ser.* 368, 492 (1988).
49. C. J. Brinker and G. W. Scherer, "Sol-Gel Science: The Physics and Chemistry of Sol-Gel Processing." Academic Press, 1990.
50. T. Hisanaga and K. Tanaka, *J. Hazard. Mater.* 93, 331 (2002).
51. D. R. Paul and D. R. Kemp, *J. Polym. Sci., Polym. Symp.* No. 41, 79 (1973).
52. S. Kulprathipanja, E. W. Funk, S. S. Kulkarni, and Y. A. Chang, U.S. Patent 4735193, 1988.
53. S. Kulprathipanja, R. W. Neuzil, and N. N. Li, U.S. Patent 4740219, 1988.
54. R. Mahajan, C. M. Zimmerman, and W. J. Koros, *ACS Symposium Series* 733, 277 (1999).
55. G. Cho, J.-S. Lee, D. T. Glatzhofer, B. M. Fung, W. L. Yuan, and E. A. O'Rear, *Adv. Mater.* 11, 497 (1999).
56. Y. S. Chun, K. Ha, Y.-J. Lee, J. S. Lee, H. S. Kim, Y. S. Park, and K. B. Yoon, *Chem. Commun.* 1846 (2002).
57. G. S. Lee, Y. J. Lee, K. Ha, and K. B. Yoon, *Tetrahedron* 56, 6965 (2000).
58. G. S. Lee, Y.-J. Lee, K. Ha, and K. B. Yoon, *Adv. Mater.* 13, 1491 (2001).
59. G. S. Lee, Y.-J. Lee, and K. B. Yoon, *J. Amer. Chem. Soc.* 123, 9769 (2001).
60. K. Ha, Y.-J. Lee, Y. S. Chun, Y. S. Park, G. S. Lee, and K. B. Yoon, *Adv. Mater.* 13, 594 (2001).
61. O. Talu, M. S. Sun, and D. B. Shah, *AIChE J.* 44, 681 (1998).
62. H. Suzuki, U.S. Patent 4,699,892, 1987.
63. S. A. I. Barri, G. J. Bratton, and T. D. V. Naylor, *Eur. Pat. Appl.* 481660, 1992.
64. E. R. Geus, M. J. Den Exter, and H. Van Bekkum, *J. Chem. Soc., Faraday Trans.* 88, 3101 (1992).
65. C. Bai, M.-D. Jia, J. L. Falconer, and R. D. Noble, *J. Membrane Sci.* 105, 79 (1995).
66. A. J. Burggraaf, Z. A. E. P. Vroon, K. Keizer, and H. Verweij, *J. Membrane Sci.* 144, 77 (1998).
67. W. O. Haag and J. G. Tsikoyiannis, U.S. Patent 5069794, 1991.
68. L. Gora, J. C. Jansen, and T. Maschmeyer, *Chem. Eur. J.* 6, 2537 (2000).
69. G. Xomeritakis, A. Gouzinis, S. Nair, T. Okubo, M. He, R. M. Overney, and M. Tsapatsis, *Chem. Eng. Sci.* 54, 3521 (1999).
70. T. Nishide, T. Hamatsu, Y. Kiyozumi, and F. Mizukami, *J. Mater. Chem.* 12, 1465 (2002).
71. Y. Yan, M. E. Davis, and G. R. Gavalas, *J. Membrane Sci.* 126, 53 (1997).
72. Y. Yan, M. E. Davis, and G. R. Gavalas, *J. Membrane Sci.* 123, 95 (1997).
73. Y. Murakami, *Stud. Surf. Sci. Catal.* 44, 177 (1989).
74. M. Nomura, T. Yamaguchi, and S. I. Nakao, *J. Membrane Sci.* 187, 203 (2001).
75. K. Aoki, V. A. Tuan, J. L. Falconer, and R. D. Noble, *Microporous Mesoporous Mater.* 39, 485 (2000).
76. V. A. Tuan, J. L. Falconer, and R. D. Noble, *Microporous Mesoporous Mater.* 41, 269 (2000).
77. X. Lin, J. L. Falconer, and R. D. Noble, *Chem. Mater.* 10, 3716 (1998).
78. E. R. Geus and H. van Bekkum, *Zeolites* 15, 333 (1995).
79. J. Dong, Y. S. Lin, M. Z. C. Hu, R. A. Peascoe, and E. A. Payzant, *Microporous Mesoporous Mater.* 34, 241 (2000).
80. M. Lassinantti, F. Jareman, J. Hedlund, D. Creaser, and J. Sterte, *Catalysis Today* 67, 109 (2001).
81. J. Hedlund, M. Noack, P. Kolsch, D. Creaser, J. Caro, and J. Sterte, *J. Membrane Sci.* 159, 263 (1999).
82. R. Lai and G. R. Gavalas, *Microporous Mesoporous Mater.* 38, 239 (2000).
83. G. Bonilla, M. Tsapatsis, D. G. Vlachos, and G. Xomeritakis, *J. Membrane Sci.* 182, 103 (2001).
84. K. H. Bennett, K. D. Cook, J. L. Falconer, and R. D. Noble, *Anal. Chem.* 71, 1016 (1999).
85. T. Q. Gardner, A. I. Flores, R. D. Noble, and J. L. Falconer, *AIChE J.* 48, 1155 (2002).
86. N. Nishiyama, M. Miyamoto, Y. Egashira, and K. Ueyama, *Chem. Commun.* 1746 (2001).
87. A. Julbe, D. Farrusseng, J. C. Jalibert, C. Mirodatos, and C. Guizard, *Catal. Today* 56, 199 (2000).
88. L. T. Y. Au, J. L. H. Chau, C. T. Ariso, and K. L. Yeung, *J. Membrane Sci.* 183, 269 (2001).
89. E. Piera, C. Tellez, J. Coronas, M. Menendez, and J. SaIntamaria, *Catal. Today* 67, 127 (2001).
90. M. A. Salomon, J. Coronas, M. Menendez, and J. Santamaria, *Applied Catalysis, A: General* 200, 201 (2000).
91. D. Casanave, P. Ciavarella, K. Fiaty, and J. A. Dalmon, *Chem. Eng. Sci.* 54, 2807 (1999).
92. G. Barbieri, G. Marigliano, G. Golemme, and E. Drioli, *Chemical Engineering Journal* 85, 53 (2002).
93. Y. Takata, T. Tsuru, T. Yoshioka, and M. Asaeda, *Microporous Mesoporous Mater.* 54, 257 (2002).
94. T. Bein, K. Brown, G. C. Frye, and C. J. Brinker, *J. Am. Chem. Soc.* 111, 7640 (1989).
95. D. J. Leary, U.S. Patent 4347732, 1982.
96. G. A. Ozin, A. Kuperman, and A. Stein, *Angewandte Chemie* 101, 373 (1989).
97. K. Alberti, J. Haas, C. Plog, and F. Fetting, *Catal. Today* 8, 509 (1991).
98. Y. Yan and T. Bein, *J. Phys. Chem.* 96, 9387 (1992).
99. K. Alberti and F. Fetting, *Sensors and Actuators, B: Chemical* 21, 39 (1994).
100. T. Bein and Y. Yan, *ACS Symposium Series* 561, 16 (1994).
101. J. C. Jansen, D. Kashchiev, and A. Erdem-Senatalar, *Stud. Surf. Sci. Catal.* 85, 215 (1994).
102. J. H. Koegler, H. W. Zandbergen, J. L. N. Harteveld, M. S. Nieuwenhuizen, J. C. Jansen, and H. van Bekkum, *Stud. Surf. Sci. Catal.* 84, 307 (1994).
103. S. Feng, Y. Yan, and T. Bein, *Stud. Surf. Sci. Catal.* 98, 281 (1995).
104. C. Plog, W. Maunz, P. Kurzweil, E. Obermeier, and C. Scheibe, *Sens. Actuators, B: Chemical* B25, 403 (1995).
105. L. J. Ball, Thesis, University of Texas, Dallas, 1996.
106. K. Fukui and S. Nishida, *Chemical Sensors* 12, 25 (1996).
107. N. M. Peachey, R. C. Dye, P. D. Ries, M. Warren, and M. Olken, *J. Porous Mat.* 2, 331 (1996).
108. K. J. Balkus, Jr., L. J. Ball, M. E. Gimon-Kinsel, J. M. Anthony, and B. E. Gnade, *Sens. Actuators, B: Chemical* B42, 67 (1997).
109. S. Mintova, B. Schoeman, V. Valtchev, J. Sterte, S. Mo, and T. Bein, *Adv. Mater.* 9, 585 (1997).
110. K. Fukui, *Jpn. Kokai Tokkyo Koho*, 09105731, 1997.
111. K. Fukui and S. Nishida, *Sens. Actuators, B: Chemical* B45, 101–106 (1997).
112. H. Tsuchiya, I. Sasaki, A. Harano, T. Okubo, and M. Sadakata, *Materials Research Society Symposium Proceedings* 454, 297 (1997).
113. C. J. Brinker, *Current Opinion in Colloid & Interface Science* 3, 166 (1998).
114. M.-K. Jung, S.-S. Hong, and M.-H. Kim, *Korean Journal of Chemical Engineering* 15, 552 (1998).
115. T. Munoz, Jr. and K. J. Balkus, Jr., *Chem. Mater.* 10, 4114 (1998).
116. F. Mizukami, *Stud. Surf. Sci. Catal.* 125, 1 (1999).
117. A. A. Shul'ga, B. K. Zuev, and V. V. Lontsov, *J. Anal. Chem.* (translation of *Zhurnal Analiticheskoi Khimii*) 54, 449 (1999).

118. W. L. Rauch and M. Liu, "Proceedings-Electrochemical Society," 2000-32, 408 (2000).
119. S. Mintova, S. Mo, and T. Bein, *Chem. Mater.* 13, 901 (2001).
120. G. Cooper, *PCT Int. Appl.* 0190732, 2001.
121. S. Mintova and T. Bein, *Microporous Mesoporous Mater.* 50, 159 (2001).
122. M. L. Hamlaoui, K. Reybier, M. Marrakchi, N. Jaffrezic-Renault, C. Martelet, R. Kherrat, and A. Walcarius, *Analytica Chimica Acta* 466, 39 (2002).
123. I. Sasaki, H. Tsuchiya, M. Nishioka, M. Sadakata, and T. Okubo, *Sens. Actuators, B: Chemical* 86, 26 (2002).
124. K. Tanaka, C.-K. Choo, S. Sumi, Y. Kamitani, T. Fujii, K. Satoh, K.-I. Fukuda, R. Nakata, M. Yoshimune, Y. Yoshinaga, and T. Okuhara, *J. Phy. Chem. B* 106, 4155 (2002).
125. A. Sands and S. M. Dick, *Vacuum* 16, 691 (1966).
126. S. P. Davis, E. V. R. Borgstedt, and S. L. Suib, *Chem. Mater.* 2, 712 (1990).
127. S. R. Dunne, M. J. McKeon, A. P. Cohen, and A. S. Behan, U.S. Patent 5120694, 1992.
128. S. R. Dunne and S. M. Taqvi, "6th Conference on Fundamentals of Adsorption," Giens, France, May 24–28, 1998, 1101.
129. M. Kamikubo, T. Ikeda, C. Hayashi, and G. Masuda, *Jpn. Kokai Tokkyo Koho* 06170172, 1994.
130. C. D. Madhusoodana, R. N. Das, Y. Kameshima, A. Yasumori, and K. Okada, *J. Porous Mater.* 8, 265 (2001).
131. C. D. Madhusoodana, R. N. Das, Y. Kameshima, A. Yasumori, and K. Okada, *Microporous Mesoporous Mater.* 46, 249 (2001).
132. Z. Shan, W. E. J. van Kooten, O. L. Oudshoorn, J. C. Jansen, H. van Bekkum, C. M. van den Bleek, and H. P. A. Calis, *Microporous Mesoporous Mater.* 34, 81 (2000).
133. D. G. Braithwaite, U.S. Patent 3697446, 1972.
134. E. W. Albers and G. C. Edwards, U.S. Patent 3730910, 1973.
135. P. D. Caesar, U.S. Patent 4692423, 1987.
136. E. V. Rebrov, G. B. F. Seijger, H. P. A. Calis, M. H. J. M. de Croon, C. M. van den Bleek, and J. C. Schouten, *Applied Catalysis, A: General* 206, 125 (2001).
137. Y. S. S. Wan, J. L. H. Chau, A. Gavriilidis, and K. L. Yeung, *Microporous Mesoporous Mater.* 42, 157 (2001).
138. Z. Wang, H. Wang, A. Mitra, L. Huang, and Y. Yan, *Adv. Mater.* 13, 746 (2001).
139. Z. Wang, H. Wang, A. Mitra, L. Huang, and Y. Yan, *Stud. Surf. Sci. Catal.* 135, 20-P-11(2001).
140. E. M. Flanigen, J. M. Bennett, R. W. Grose, J. P. Cohen, R. L. Patton, R. M. Kirchner, and J. V. Smith, *Nature* 271, 512 (1978).
141. A. Mitra, Z. Wang, T. Cao, H. Wang, L. Huang, and Y. Yan, *J. Electrochem. Soc.* 149, B472 (2002).
142. A. Mitra, C. W. Kirby, Z. Wang, L. Huang, H. Wang, Y. Huang, and Y. Yan, *Microporous Mesoporous Mater.* 54, 175 (2002).
143. H. Wang, Z. Wang, X. Cheng, A. Mitra, L. Huang, and Y. Yan, *Stud. Surf. Sci. Catal.* 135, 22-O-04 (2001).
144. J. C. Fischer, Jr., U.S. Patent 4769053, 1988.
145. A. Erdem-Senatalar, M. Tatlier, and M. Urgen, *Microporous Mesoporous Mater.* 32, 331 (1999).
146. M. Tatlier and A. Erdem-Senatalar, *Stud. Surf. Sci. Catal.* 125, 101 (1999).
147. H. W. Hillhouse and M. T. Tuominen, *Microporous Mesoporous Mater.* 47, 39 (2001).
148. J. Dong, K. Wegner, and Y. S. Lin, *J. Membrane Sci.* 148, 233 (1998).
149. L. Gora, N. Nishiyama, J. C. Jansen, F. Kapteijn, V. Teplyakov, and T. Maschmeyer, *Separation and Purification Technology* 22–23, 223 (2001).
150. P. Meriaudeau, A. Thangaraj, and C. Naccache, *Microporous Materials* 4, 213 (1995).
151. R. Lai, Y. Yan, and G. R. Gavalas, *Microporous Mesoporous Mater.* 37, 9 (2000).
152. K. Kusakabe, S. Yoneshige, A. Murata, and S. Morooka, *J. Membrane Sci.* 116, 39 (1996).
153. C. L. Flanders, V. A. Tuan, R. D. Noble, and J. L. Falconer, *J. Membrane Sci.* 176, 43 (2000).
154. B. Oonkhanond and M. E. Mullins, *J. Membrane Sci.* 194, 3 (2001).
155. J. L. Hang Chau, C. Tellez, K. L. Yeung, and K. Ho, *J. Membrane Sci.* 164, 257 (2000).
156. E. R. Geus, H. van Bekkum, W. J. W. Bakker, and J. A. Moulijin, *Microporous Mater.* 1, 131 (1993).
157. B. S. Kang and G. R. Gavalas, *Ind. & Eng. Chem. Res.* 41, 3145 (2002).
158. R. Lai and G. R. Gavalas, *Ind. & Eng. Chem. Res.* 37, 4275 (1998).
159. S. M. Holmes, C. Markert, R. J. Plaisted, J. O. Forrest, J. R. Agger, M. W. Anderson, C. S. Cundy, and J. Dwyer, *Chem. Mater.* 11, 3329 (1999).
160. M. C. Lovallo, A. Gouzinis, and M. Tsapatsis, *AIChE J.* 44, 1903 (1998).
161. M. Noack, P. Kolsch, J. Caro, M. Schneider, P. Toussaint, and I. Sieber, *Microporous Mesoporous Mater.* 35–36, 253 (2000).
162. G. Bonilla, D. G. Vlachos, and M. Tsapatsis, *Microporous Mesoporous Mater.* 42, 191 (2001).
163. C. S. Tsay and A. S. T. Chiang, *AIChE J.* 46, 616 (2000).
164. S. Alfaro, M. Arruebo, J. Coronas, M. Menendez, and J. Santamaria, *Microporous Mesoporous Mater.* 50, 195 (2001).
165. T. Sano, M. Hasegawa, Y. Kawakami, and H. Yanagishita, *J. Membrane Sci.* 107, 193 (1995).
166. S. Li, V. A. Tuan, J. L. Falconer, and R. D. Noble, *Chem. Mater.* 13, 1865 (2001).
167. M. Nomura, T. Bin, and S.-i. Nakao, *Sep. Purif. Technol.* 27, 59 (2002).
168. T. Sano, H. Yanagishita, Y. Kiyozumi, D. Kitamoto, and F. Mizukami, *Chem. Lett.* 2413 (1992).
169. Q. Liu, R. D. Noble, J. L. Falconer, and H. H. Funke, *J. Membrane Sci.* 117, 163 (1996).
170. K. Wegner, J. Dong, and Y. S. Lin, *J. Membrane Sci.* 158, 17–27 (1999).
171. J. Hedlund, J. Sterte, M. Anthonis, A.-J. Bons, B. Carstensen, N. Corcoran, D. Cox, H. Deckman, W. De Gijnst, P.-P. de Moor, F. Lai, J. McHenry, W. Mortier, J. Reinoso, and J. Peters, *Microporous Mesoporous Mater.* 52, 179 (2002).
172. H. H. Funke, M. G. Kovalchick, J. L. Falconer, and R. D. Noble, *Ind. & Eng. Chem. Res.* 35, 1575 (1996).
173. J. Coronas, R. D. Noble, and J. L. Falconer, *Ind. & Eng. Chem. Res.* 37, 166 (1998).
174. C. J. Gump, R. D. Noble, and J. L. Falconer, *Ind. & Eng. Chem. Res.* 38, 2775 (1999).
175. M. P. Bernal, M. Bardaji, J. Coronas, and J. Santamaria, *J. Membrane Sci.* 203, 209 (2002).
176. J. C. Poshusta, R. D. Noble, and J. L. Falconer, *J. Membrane Sci.* 160, 115 (1999).
177. L. J. P. Van Den Broeke, F. Kapteijn, and J. A. Moulijn, *Chem. Eng. Sci.* 54, 259 (1998).
178. E. Piera, C. A. M. Brenninkmeijer, J. Santamaria, and J. Coronas, *J. Membrane Sci.* 201, 229 (2002).
179. Y. S. Cheng, M. A. Pena, J. L. Fierro, D. C. W. Hui, and K. L. Yeung, *J. Membrane Sci.* 204, 329 (2002).
180. J. Dong, Y. S. Lin, and W. Liu, *AIChE J.* 46, 1957 (2000).
181. L. J. P. Van Den Broeke, W. J. W. Bakker, F. Kapteijn, and J. A. Moulijn, *Chem. Eng. Sci.* 54, 245 (1998).
182. W. J. W. Bakker, F. Kapteijn, J. Poppe, and J. A. Moulijn, *J. Membrane Sci.* 117, 57 (1996).
183. C. J. Gump, X. Lin, J. L. Falconer, and R. D. Noble, *J. Membrane Sci.* 173, 35 (2000).
184. L. T. Y. Au and K. L. Yeung, *J. Membrane Sci.* 194, 33 (2001).
185. J. M. van de Graaf, F. Kapteijn, and J. A. Moulijn, *J. Membrane Sci.* 144, 87 (1998).

186. J. M. Van de Graaf, E. Van der Bijl, A. Stol, F. Kapteijn, and J. A. Moulijn, *Ind. & Eng. Chem. Res.* 37, 4071 (1998).
187. J. M. van de Graaf, F. Kapteijn, and J. A. Moulijn, *Microporous Mesoporous Mater.* 35–36, 267 (2000).
188. A. J. Burggraaf, *J. Membrane Sci.* 155, 45 (1999).
189. J. Romero, S. Le Cam, J. Sanchez, A. Saavedra, and G. M. Rios, *Chem. Eng. Sci.* 56, 3139 (2001).
190. S. Farooq and I. A. Karimi, *J. Membrane Sci.* 186, 109 (2001).
191. B. Millot, A. Methivier, H. Jobic, H. Moueddeb, and J. A. Dalmon, *Microporous Mesoporous Mater.* 38, 85 (2000).
192. R. Krishna, *International Communications in Heat and Mass Transfer* 28, 337 (2001).
193. R. Krishna and D. Paschek, *Ind. & Eng. Chem. Res.* 39, 2618 (2000).
194. N. Nishiyama, L. Gora, V. Teplyakov, F. Kapteijn, and J. A. Moulijn, *Sep. Puri. Technol.* 22–23, 295 (2001).
195. M. Nomura, T. Yamaguchi, and S.-i. Nakao, *J. Membrane Sci.* 144, 161 (1998).
196. S. Suzuki, H. Takaba, T. Yamaguchi, and S.-I. Nakao, *J. Phys. Chem. B* 104, 1971 (2000).
197. M. G. Martin, A. P. Thompson, and T. M. Nenoff, *J. Chem. Phys.* 114, 7174 (2001).
198. H. Takaba, A. Koyama, and S.-I. Nakao, *J. Phys. Chem. B* 104, 6353 (2000).
199. P. I. Pohl, G. S. Heffelfinger, and D. M. Smith, *Mol. Phys.* 89, 1725 (1996).
200. P. H. Nelson, M. Tsapatsis, and S. M. Auerbach, *J. Membrane Sci.* 184, 245 (2001).
201. E. I. Basaldella, A. Kikot, J. F. Bengoa, and J. C. Tara, *Materials Letters* 52, 350 (2002).
202. R. van der Vaart, H. Bosch, and T. Reith, *Microporous Mater.* 9, 203 (1997).
203. T. Sano, Y. Kiyozumi, F. Mizukami, H. Takaya, T. Mouri, and M. Watanabe, *Zeolites* 12, 131 (1992).
204. T. Sano, Y. Kiyozumi, M. Kawamura, F. Mizukami, H. Takaya, T. Mouri, W. Inaoka, Y. Toida, M. Watanabe, and K. Toyoda, *Zeolites* 11, 842 (1991).
205. M.-K. Jung, M.-H. Kim, and S.-S. Hong, *Microporous Mesoporous Mater.* 26, 153 (1998).
206. J. H. Koegler, H. Van Bekkum, and J. C. Jansen, *Zeolites* 19, 262 (1997).
207. H. H. Xu, D. B. Shah, and O. Talu, *Zeolites* 19, 114 (1997).
208. A. Iwasaki, T. Sano, and Y. Kiyozumi, *Microporous Mesoporous Mater.* 38, 75 (2000).
209. V. Valtchev, J. Hedlund, B. J. Schoeman, J. Sterte, and S. Mintova, *Microporous Mater.* 8, 93 (1997).
210. V. Engstrom, B. Mihailova, J. Hedlund, A. Holmgren, and J. Sterte, *Microporous Mesoporous Mater.* 38, 51 (2000).
211. J. Sterte, S. Mintova, G. Zhang, and B. J. Schoeman, *Zeolites* 18, 387 (1997).
212. A. Gouzinis and M. Tsapatsis, *Chem. Mater.* 10, 2497 (1998).
213. K. Morawetz, J. Reiche, H. Kamusewitz, H. Kosmella, R. Ries, M. Noack, and L. Brehmer, *Colloids and Surfaces, A: Physicochemical and Engineering Aspects* 198–200, 409 (2002).
214. D. Coutinho and K. J. Balkus, *Microporous Mesoporous Mater.* 52, 79 (2002).
215. K. Weh, M. Noack, K. Hoffmann, K. P. Schroder, and J. Caro, *Microporous Mesoporous Mater.* 54, 15 (2002).
216. Y. Hasegawa, K. Watanabe, K. Kusakabe, and S. Morooka, *Sep. Purif. Technol.* 22–23, 319 (2001).
217. Y. Hasegawa, K.-I. Sotowa, K. Kusakabe, and S. Morooka, *Microporous Mesoporous Mater.* 53, 37 (2002).
218. K. Kusakabe, T. Kuroda, and S. Morooka, *J. Membrane Sci.* 148, 13 (1998).
219. A. Chatterjee and T. Iwasaki, *J. Phys. Chem. A* 103, 9857 (1999).
220. K. Mizukami, H. Takaba, Y. Kobayashi, Y. Oumi, R. V. Belosludov, S. Takami, M. Kubo, and A. Miyamoto, *J. Membrane Sci.* 188, 21 (2001).
221. G. Clet, J. C. Jansen, and H. van Bekkum, *Chem. Mater.* 11, 1696 (1999).
222. K. Weh, M. Noack, I. Sieber, and J. Caro, *Microporous Mesoporous Mater.* 54, 27 (2002).
223. V. Nikolakis, G. Xomeritakis, A. Abibi, M. Dickson, M. Tsapatsis, and D. G. Vlachos, *J. Membrane Sci.* 184, 209 (2001).
224. A. M. Goossens, B. H. Wouters, V. Buschmann, and J. A. Martens, *Adv. Mater.* 11, 561 (1999).
225. A. M. Goossens, B. H. Wouters, P. J. Grobet, V. Buschmann, L. Fiermans, and J. A. Martens, *European Journal of Inorganic Chemistry* 1167 (2001).
226. G. J. Myatt, P. M. Budd, C. Price, and S. W. Carr, *J. Mater. Chem.* 2, 1103 (1992).
227. S. M. Holmes, M. Schmitt, C. Markert, R. J. Plaisted, J. O. Forrest, P. N. Sharratt, A. A. Garforth, C. S. Cundy, and J. Dwyer, *Chemical Engineering Research and Design* 78, 1084 (2000).
228. D. Shah, K. Kissick, A. Ghorpade, R. Hannah, and D. Bhattacharyya, *J. Membrane Sci.* 179, 185 (2000).
229. K.-I. Okamoto, H. Kita, K. Horii, and K. Tanaka, *Ind. & Eng. Chem. Res.* 40, 163 (2001).
230. M. Kondo, M. Komori, H. Kita, and K.-I. Okamoto, *J. Membrane Sci.* 133, 133 (1997).
231. X. Xu, W. Yang, J. Liu, and L. Lin, *Microporous Mesoporous Mater.* 43, 299 (2001).
232. X. Xu, W. Yang, J. Liu, X. Chen, L. Lin, N. Stroh, and H. Brunner, *Chem. Commun.* 603 (2000).
233. L. C. Boudreau and M. Tsapatsis, *Chem. Mater.* 9, 1705 (1997).
234. L. C. Boudreau, J. A. Kuck, and M. Tsapatsis, *J. Membrane Sci.* 152, 41 (1999).
235. S. Kornic and M. Baker, *Chem. Commun.* 1700 (2002).
236. K. Aoki, K. Kusakabe, and S. Morooka, *AIChE J.* 46, 221 (2000).
237. Y. Yan and T. Bein, *J. Amer. Chem. Soc.* 117, 9990 (1995).
238. I. Kumakiri, T. Yamaguchi, and S.-I. Nakao, *J. Chem. Eng. Jpn.* 33, 333 (2000).
239. J. Lin and S. Murad, *Mol. Phys.* 99, 1175 (2001).
240. S. Mintova, V. Valtchev, and L. Konstantinov, *Zeolites* 17, 462 (1996).
241. T. Sano, Y. Kiyozumi, K. Maeda, M. Toba, S. Niwa, F. Mizukami, I. Mukoyoshi, and H. Shoji, *J. Mol. Catal.* 77, L19 (1992).
242. T. Sano, Y. Kiyozumi, K. Maeda, M. Toba, S. Niwa, and F. Mizukami, *Proc. 9th Int. Zeolite Conf.* 1, 239 (1993).
243. T.-G. Tsai, H.-C. Shih, S.-J. Liao, and K.-J. Chao, *Microporous Mesoporous Mater.* 22, 333 (1998).
244. C. J. Gump, V. A. Tuan, R. D. Noble, and J. L. Falconer, *Ind. & Eng. Chem. Res.* 40, 565 (2001).
245. B. K. Adnadjevic, J. D. Jovanovic, and S. P. Gajinov, *Hemijska Industrija* 50, 537 (1996).
246. B. Adnadjevic, J. Jovanovic, and S. Gajinov, *J. Membrane Sci.* 136, 173 (1997).
247. I. Girnus, M. M. Pohl, J. Richter-Mendau, M. Schneider, M. Noack, D. Venzke, and J. Caro, *Adv. Mater.* 7, 711 (1995).
248. Y. H. Chiou, T. G. Tsai, H. C. Shih, C. N. Wu, and K. J. Chao, *J. Mater. Sci. Lett.* 15, 952 (1996).
249. K. J. Balkus, Jr., S. J. Riley, and B. E. Gnade, *Materials Research Society Symposium Proceedings* 351, 437 (1994).
250. S. Mintova, S. Mo, and T. Bein, *Chem. Mater.* 10, 4030–4036 (1998).
251. S. Feng and T. Bein, *Science* 265, 1839 (1994).
252. T. G. Tsai, K. J. Chao, X. J. Guo, S. L. Sung, C. N. Wu, Y. L. Wang, and H. C. Shih, *Adv. Mater.* 9, 1154 (1997).
253. C. N. Wu, K. J. Chao, T. G. Tsai, Y. H. Chiou, and H. C. Shih, *Adv. Mater.* 8, 1008 (1996).

254. K.-J. Chao, C.-N. Wu, H.-C. Shih, and T.-G. Tsai, U.S. Patent 6060145, 2000.
255. M. Noack, P. Koelsch, D. Venzke, P. Toussaint, and J. Caro, *Microporous Mater.* 3, 201 (1994).
256. M. Noack, P. Koelsch, D. Venzke, P. Toussaint, and J. Caro, *Stud. Surf. Sci. Catal.* 98, 276 (1995).
257. D. Farrusseng, A. Julbe, and C. Guizard, *Sep. Purif. Technol.* 25, 137 (2001).
258. J. Caro, F. Marlow, K. Hoffmann, J. Kornatowski, I. Girnus, M. Noack, and P. Koelsch, "Proceedings—Polish-German Zeolite Colloquium," 2nd, Torun, Pol., April 20–22, 1995, 186 (1995).
259. D. S. Sholl, *Chemical Engineering Journal* 74, 25 (1999).
260. D. S. Sholl, *Ind. & Eng. Chem. Res.* 39, 3737 (2000).
261. H. Mimura, T. Tezuka, and K. Akiba, *J. Nucl. Sci. Technol.* 32, 1250 (1995).
262. H. Mimuara, T. Tezuka, and K. Akiba, *J. Nucl. Sci. Technol* 33, 892 (1996).
263. N. Nishiyama, K. Ueyama, and M. Matsukata, *Microporous Mater.* 7, 299 (1996).
264. V. A. Tuan, S. Li, J. L. Falconer, and R. D. Noble, *Chem. Mater.* 14, 489 (2002).
265. X. Li and S. Xiang, *Stud. Surf. Sci. Catal.* 135, 20-P-12 (2001).
266. B. J. Schoeman, E. Babouchkina, S. Mintova, V. P. Valtchev, and J. Sterte, *J. Porous Mater.* 8, 13 (2001).
267. O. L. Oudshoorn, M. Janissen, W. E. J. Van Kooten, J. C. Jansen, H. Van Bekkum, C. M. Van den Bleek, and H. P. A. Calis, *Chem. Eng. Sci.* 54, 1413 (1999).
268. Y. Li, S. Yu, X. Yuan, and H. Wang, *Ind. & Eng. Chem. Res.* 41, 4936 (2002).
269. V. Valtchev and S. Sferdjella, *Stud. Surf. Sci. Catal.* 135, 3371 (2001).
270. H. Lee and P. K. Dutta, *Microporous Mesoporous Mater.* 38, 151 (2000).
271. E. Piera, M. A. Salomon, J. Coronas, M. Menendez, and J. Santamaria, *J. Membrane Sci.* 149, 99 (1998).
272. M. A. Salomon, J. Coronas, M. Menendez, and J. Santamaria, *Chem. Commun.* 125 (1998).
273. E. Piera, M. P. Bernal, M. A. Salomon, J. Coronas, M. Menendez, and J. Santamaria, *Stud. Surf. Sci. Catal.* 125, 189 (1999).
274. S. M. J. Zaidi, S. D. Mikhailenko, and S. Kaliaguine, "Proceedings of the Intersociety Energy Conversion Engineering Conference," 1998.
275. T. Wakihara, J. Plevert, S. Nair, M. Tsapatsis, S. Yamakita, Y. Ogawa, H. Komiyama, M. Yoshimura, M. E. Davis, and T. Okubo, *Stud. Surf. Sci. Catal.* 135, 02-P-28 (2001).
276. L. Zhang and H. Wang, *J. Mater. Sci.* 37, 1491 (2002).
277. J. C. Poshusta, V. A. Tuan, J. L. Falconer, and R. D. Noble, *Ind. & Eng. Chem. Res.* 37, 3924 (1998).
278. X. Li, W. Geng, J. Xu, Y. Fu, and L. Wang, *Huagong Xuebao* 52, 944 (2001).
279. J. C. Poshusta, V. A. Tuan, E. A. Pape, R. D. Noble, and J. L. Falconer, *AIChE J.* 46, 779 (2000).
280. J. C. Poshusta, R. D. Noble, and J. L. Falconer, *J. Membrane Sci.* 186, 25 (2001).
281. K. J. Balkus, Jr., T. Munoz, and M. E. Gimon-Kinsel, *Chem Mater.* 10, 464 (1998).
282. K. J. Balkus and A. S. Scott, *Microporous Mesoporous Mater.* 34, 31 (2000).
283. K. Tanaka, R. Yoshikawa, C. Ying, H. Kita, and K. I. Okamoto, *Catal. Today* 67, 121 (2001).
284. K. Tanaka, R. Yoshikawa, C. Ying, H. Kita, and K.-I. Okamoto, *Chem. Eng. Sci.* 57, 1577 (2002).
285. M. Matsukata, N. Nishiyama, and K. Ueyama, *Microporous Mater.* 7, 109 (1996).
286. N. Nishiyama, T. Matsufuji, K. Ueyama, and M. Masahiko, *Microporous Mater.* 12, 293 (1997).
287. T. Matsufuji, N. Nishiyama, K. Ueyama, and M. Matsukata, *Microporous Mesoporous Mater.* 32, 159 (1999).
288. T. Matsufuji, S. Nakagawa, N. Nishiyama, M. Matsukata, and K. Ueyama, *Microporous Mesoporous Mater.* 38, 43 (2000).
289. J. Dong and Y. S. Lin, *Ind. & Eng. Chem. Res.* 37, 2404 (1998).
290. M. P. Bernal, J. Coronas, M. Menendez, and J. Santamaria, *Ind. & Eng. Chem. Res.* 41, 5071 (2002).
291. M. P. Bernal, E. Piera, J. Coronas, M. Menendez, and J. Santamaria, *Catal. Today* 56, 221 (2000).
292. X. Lin, E. Kikuchi, and M. Matsukata, *Chem. Commun.* 957 (2000).
293. S. Yamazaki and K. Tsutsumi, *Microporous Mater.* 5, 245 (1995).
294. S. Yamazaki and K. Tsutsumi, *Adsorption* 3, 165 (1997).
295. G. Li, X. Lin, E. Kikuchi, and M. Matsukata, *Stud. Surf. Sci. Catal.* 135, 3153 (2001).
296. N. Nishiyama, K. Ueyama, and M. Matsukata, *AIChE J.* 43, 2724 (1997).
297. M. Matsukata, U.S. Pat. Appl. 20010012505, 2001.
298. N. M. Peachey, R. C. Dye, and P. D. Ries, "Proceedings of the International Conference on Lasers," 666 (1995).
299. J. Li, J. Dong, G. Liu, Y. Shi, J. Cao, W. Xu, and F. Wu, *Reaction Kinetics and Catalysis Letters* 47, 287 (1992).
300. H. Kanesaka, N. Kachi, and G. Masuda, *Jpn. Kokai Tokkyo Koho* 05285392, 1993.
301. G. Johansson, L. Risinger, and L. Faelth, *Anal. Chim. Acta* 119, 25 (1980).
302. J. Kjaer, S. Yde-Andersen, N. A. Knudsen, and E. Skou, *Solid State Ionics* 46, 169 (1991).
303. N. Rao, T. P. Andersen, and P. Ge, *Solid State Ionics* 72, 334 (1994).
304. M. Cheng, D. Tan, and X. Bao, *Chem. Commun.* 1713 (2000).
305. M. Cheng, G. Hu, D. Tan, and X. Bao, *Microporous Mesoporous Mater.* 50, 69 (2001).
306. K. J. Balkus, G. Gbery, and Z. Deng, *Microporous Mesoporous Mater.* 52, 141 (2002).
307. C. M. Braunbarth, L. C. Boudreau, and M. Tsapatsis, *J. Membrane Sci.* 174, 31 (2000).
308. M. J. Den Exter, H. van Bekkum, C. J. M. Rijn, F. Kapteijn, J. A. Moulijn, H. Schellevis, and C. I. N. Beenakker, *Zeolites* 19, 13 (1997).
309. Y. Yan, S. R. Chaudhuri, and A. Sarkar, *Chem. Mater.* 8, 473 (1996).
310. T. H. Metzger, S. Mintova, and T. Bein, *Microporous Mesoporous Mater.* 43, 191 (2001).
311. L. Huang, Z. Wang, J. Sun, L. Miao, Q. Li, Y. Yan, and D. Zhao, *J. Am. Chem. Soc.* 122, 3530 (2000).
312. L. Huang, Z. Wang, H. Wang, J. Sun, Q. Li, D. Zhao, and Y. Yan, *Microporous Mesoporous Mater.* 48, 73 (2001).
313. K. T. Jung and Y. G. Shul, *Chem. Mater.* 9, 420 (1997).
314. K. T. Jung and Y. G. Shul, *J. Membrane Sci.* 191, 189 (2001).
315. H. Wang, L. Huang, Z. Wang, A. Mitra, and Y. Yan, *Chem. Commun.* 1364 (2001).
316. V. Valtchev, B. J. Schoeman, J. Hedlund, S. Mintova, and J. Sterte, *Zeolites* 17, 408 (1996).
317. K. T. Jung and Y. G. Shul, *Microporous Mesoporous Mater.* 21, 281 (1998).
318. H. Wang, L. Huang, A. Holmberg Brett, and Y. Yan, *Chem. Commun.* 1708 (2002).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanostructures Created by Lasers

E. G. Gamaly, A. V. Rode

Australian National University, Canberra, Australia

CONTENTS

1. INTRODUCTION

Tiny edifices with a size of around 1–10 nm (1 nm $= 10^{-9}$ m) are called nanocrystals or nanoclusters. These structures are aggregates of atoms or molecules containing from a few atoms to a few thousands atoms. The novelty of these systems relates to the fact that they have an atomic arrangement and material properties drastically different from the same material in bulk. Nanoscience aims at the study of electronic, material, optical, and structural properties of tiny clusters and their composites. The goal of nanotechnology is the production of new and different things—smaller and with new properties. It is emerging as the next technological revolution, promising great advances in many fields, from information technology and medicine to artificial intelligence and environmental restoration.

Richard Feynman in his prophetic lecture "Plenty of Room at the Bottom" in 1959 expressed a strong belief that someday we would be able to manipulate matter on the atomic scale: "The principles of physics, as far as I can see, do not speak against the possibility of maneuvering things atom by atom. It is not an attempt to violate any laws; it is something, in principle, that can be done ...." Thus, the core idea and the ultimate goal of nanotechnology is a realization of "the possibility of manoeuvring things atom by atom": constructing a new nanomaterial by placing atom by atom at the proper construction site at a proper time. We are still far from realization of this idea in practice; however, many important steps in this direction have been made.

The major problem in nanotechnology lies in the development of efficient methods for the formation of nanoclusters with predictable and controlled size and properties that can then be scaled to industrial production. The primary goal in any method for nanocluster formation is to decompose the initial material into a preferably atomic state via evaporation or ablation and form a beam (or cloud) of hot atoms. The interaction of this beam with the flow of another gas (a noble gas or a reactive one) in a proper temperature-controlled environment creates the temperature and density conditions for the formation of a new structure. The hot atoms from the beam collide with the ambient gas atoms, cool down, and attach to each other, binding and forming a new structure in a self-assembly mode. The newly created material should then be annealed appropriately in order to transform it into a phase state which is stable at room temperature and atmospheric pressure.

The formation of different structures in a self-assembly fashion occurs in a variety of natural systems. In the methods for nanocluster formation, self-assembly is achieved by the proper choice of reacting materials so that they naturally bond with each other in the desired configuration. In theory, one can prepare an atomic mixture in an experimental chamber under carefully controlled conditions and allow the components to assemble themselves. This is the so-called bottom-up approach, or atom-to-atom attachment. In reality the processes of atomic beam formation, cluster creation, and annealing are fast and transient. These processes are very difficult to control and register *in-situ* even with modern diagnostic methods.

A variety of methods have already been proposed and employed for the production of nanoclusters in the past 2

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (783–809)

decades, including the use of supersonic nozzle sources, ion beam implantation, methods of colloid chemistry, chemical vapor deposition (CVD), arc discharge, ion sputtering, laser ablation, and others [1–3]. Methods for the creation of supersonic molecular and atomic beams have allowed one the production of clusters in the gas phase from different materials, including the most refractory and complex systems (fullerenes, carbon and boron nitride nanotubes, metallic and covalent clusters, complex clusters). Nanofilms of different materials and thin layers of matrix-embedded cluster composites have also been successfully deposited and characterized [1–3].

In fact, all methods for nanocluster formation stem from the original experiment for the preparation of so-called "blacks," the historical predecessor of nanocrystal networks, conceived and implemented by the American physicist Pfund in 1930 [4, 5]. He evaporated bismuth in an imperfect vacuum. The low-density gas fill acts as a confinement and a heat sink for evaporated the atoms. The hot atoms from the evaporation source cool down through collisions with the gas fill and then form tiny droplets when the supersaturated state is achieved. The process comprises three stages: nucleation, growth, and cluster network formation on a substrate [5]. The differences between all of the modern methods for nanocluster formation relate mainly to how the atomic beam is prepared. In this chapter we describe the laser ablation method for the formation and control of an atomic plume which is then used for the formation of various nanostructures under different experimental conditions.

The laser ablation process has a number of remarkable features. First, laser ablation has the highest instantaneous ablation and deposition rate among all known deposition methods, such as electron-beam deposition, magnetron sputtering, and chemical vapor deposition. Second, a laser beam is a "clean" tool: it provides direct energy transfer into the ablated material, and thus there is no source of contamination in the process. Consequently, the laser deposition process maintains the same stoichiometry in the deposited film or in a nanocluster as that in the ablated target. Third, the laser ablation process can operate over a wide range of gas pressures. Finally, it is possible to manipulate the atomic beam created by a short laser pulse at a time scale of picoseconds and a space scale of micrometers.

During the past decade, a number of complicated structures have been successfully produced by laser ablation and deposition. Examples include high-temperature superconducting films [6], various types of nitrides [7–9], complex multicomponent oxides [10, 11], nanocrystalline quantum dots [12], ferroelectric thin films [13], planar lasers, and nonlinear waveguides [14, 15]. Throughout these studies lasers of various wavelengths (from 193 nm to 10.6 μm), pulse durations (from tens of picoseconds to hundreds of nanoseconds), and intensities have been used.

Laser vaporization has been used for the production of nanocluster production since the early seventies, when commercial lasers became available [1, 2, 6]. Metal clusters [16–18], fullerenes [19], carbon and boron nitride nanotubes [20–22], silicon clusters, and many other structures have been produced with laser ablation. In all these studies, commercially available lasers were used as a heating source rather than exploring all the opportunities that a laser offers for controlling beam composition and properties. Extensive studies of laser–matter interaction, including ablation and the properties of a laser-created plume, were performed within laser fusion projects in major laser laboratories around the world starting in the early sixties [23–26]. Recently it has been understood that many results obtained in laser fusion programs can be usefully employed in laser-assisted methods for nanocluster formation in materials science research.

Laser-assisted nanostructure formation is conventionally performed in a special chamber in which an intense laser beam is focused onto a target from which the desired nanostructure should be formed. Initially, commercial Nd:YAG (neodymium doped yttrium–aluminium garnet) lasers and excimer lasers with pulse durations of 10–30 ns, repetition rates of 10–30 Hz, and average intensities per pulse of $2–4 \times 10^9$ W/cm^2 have been employed for target evaporation [16–18]. Laser evaporation generates a plume with a rather high ion temperature of a few electron volts during the pulse time. However, due to the large time gap between successive pulses, the laser plume cools down below the temperature which is necessary for cluster formation. Therefore, additional heating of the ambient gas was necessary in order to maintain the proper conditions for cluster formation. The target was placed into a furnace with a controlled temperature and with a continuous noble gas flow. With a long (nanoseconds)-pulse, low repetition rate (several Hertz) laser, it is possible to produce a pulsed flow of evaporated vapors with low time-averaged flux intensity.

The use of dual-pulsed Nd:YAG laser vaporization has proved to increase the efficiency of carbon nanotube formation [27]. Furthermore, the use of continuous-wave CO_2 lasers at an intensity level $\sim 10^5$ W/cm^2 [28] or an almost continuous 20-ms pulse [29, 30] for evaporation of carbon targets made it possible to obtain a continuous supply of hot carbon atoms into an argon-filled experimental chamber. The collision of hot carbon atoms with cold argon atoms maintains a temperature sufficient for carbon nanotube formation without additional heating in a furnace.

The next step in improving the average vapor beam intensity and control over the density and temperature of the laser-ablated vapor relates to the use of short (a few femtoseconds), intense (average intensity per pulse, $\sim 10^{13}$–10^{14} W/cm^2), low-energy (μJ) pulses delivered at high repetition rates (100 kHz–100 MHz) [31–33]. The high average intensity per pulse allows a significant number of particles, $\sim 10^{11}$ atoms per 100-fs pulse, to be removed during the short pulse time, while the high repetition rate maintains the average atomic beam intensity at a sufficiently high level of 10^{19} atoms/s.

Self-consistent choice of the optimum combination of laser parameters (pulse duration, energy, wavelength, and intensity distribution during the pulse time and in space over the focal spot) along with the target parameters allows one to produce an atomized flow of ablated vapors from any material. The temperature, density, atomic content, and degree of ionization in a cloud of vapors depend on the above combination and therefore can also be controlled.

Ablation with high repetition rate lasers also increases the average temperature of the ablated atomic beam up to

20 times compared with conventional long-pulse low repetition rate lasers. Thus additional heating during cluster formation can be eliminated and the experimental setup is simplified. The laser beam is continuously scanned over the target surface in order to prevent crater formation and to maintain a constant ablation rate for successive pulses. Fast scanning across a complex target makes possible the coevaporation of different materials. As a result, combinatorial material synthesis in the plume can be implemented. Vapor composition can be varied continuously by adjusting the dwell time on different parts of a composite target.

The laser-ablated atomic beam can be used for the formation of new structures in two fundamentally different experimental situations. In the first operation mode, the ablated plume is used for the deposition of thin films on different substrates in an evacuated chamber. The substrate to be coated is placed in the path of the laser-produced plume; the vapor adheres to the surface to form a nanofilm. In the second operation mode the ablated atomic flow interacts with a different gas fill in a chamber at different pressures, creating nanoclusters and nanocluster networks. The characteristics of the ablated plume are sensitive to a combination of laser and target parameters. Therefore, this combination should be different for the two operation modes mentioned above.

Let us consider thin film deposition by laser ablation. The main drawback to widespread industrial application of lasers for deposition of homogeneous thin films has been the formation of droplets (0.1–1 μm) in the laser plume, which spoil the surface finish of the deposited film. Many applications require the density of micrometer-sized particles to be less than one particle per cm^2 on the surface of the deposited film. The origin of these particles (droplets) relates to inhomogeneities in the target, spatial and time distribution and fluctuations of laser intensity, and the level of absorbed laser energy. While there have been attempts to prevent particles from the target from reaching the substrate using some form of mechanical filtering, there does not appear to have been any satisfactory or universal solution to the problem.

A solution to this problem has been found through deeper insight into the physics of the laser ablation process. It was proposed to use for ablation picosecond or femtosecond laser pulses delivered on a target with a repetition rate of several tens of megaHertz [31, 32]. The method was coined "ultrafast laser ablation" due to the combination of short laser pulses with a very high repetition rate. A single short low-energy pulse evaporates very few hot atoms per pulse, thereby inhibiting the condensation of droplets during fast nonequilibrium expansion. To compensate for the reduced ablated mass per pulse, high pulse repetition rates are then used to achieve a high averaged deposition rate. The application of the ultrafast laser ablation (UFLA) method for deposition of thin films has resulted in atomic surface quality films of diamondlike carbon [31, 32], with the total elimination of macroscopic particles from the film surface. Recently As_2S_3 chalcogenide optical films [34] have been produced with similar surface quality and high-volume homogeneity.

Complex nanoclustered networks of carbons were formed by evaporating a glassy carbon target in an argon-filled chamber. It was discovered that with the appropriate adjustment of the parameters of ablated flow and pressure of the ambient gas, the formation of new structures of a known material with unexpected properties results. For example, a new form of carbon, a cluster-assembled paramagnetic carbon nanofoam, was produced at an argon pressure above 0.1 Torr [37]. Conversely, at an argon pressure of 300–500 Torr, carbon nanotubes can be formed [28–30].

Recent developments clearly showed that a better understanding of basic physics and chemistry at all stages of the nanocluster formation process allowed the implementation of new methods for the prediction and control of this process. Using laser ablation, an atomized flow of any material with a controlled and predictable ablation rate, density, and temperature can be produced. The minimum number of particles ablated by a short-pulse (100 fs) laser is now reduced to $\sim 10^{11}$ atoms. This is comparable to the number of atoms in a cubic micrometer of a conventional solid. This is an important step toward the ultimate atom-to-atom attachment mode of nanocluster formation.

The research field of nanoscience and nanotechnology is a multidisciplinary one, in which the different areas of physics and chemistry overlap. In this review, we try to present a unified view of nanocluster formation by lasers, combining experimental results, theory, and simple models. The references were carefully selected to be guides in a vast sea of published papers, with emphasis on the most important topics and achievements. Therefore, the reference list is by no means exhaustive and reflects the views of the authors.

The review is organized as follows. First, we describe the lasers available for the ablation of different materials, the types of experimental installations for nanocluster production, and *in-situ* and *ex situ* diagnostic methods for observing cluster formation conditions and their structural and electronic properties (Section 2). Then we give a brief account of laser–matter interaction physics, with emphasis on the simple scaling that allows the prediction and interpretation of the experimentally observed ablation rate and plume parameters (Section 3). The processes of thin amorphous film formation and composite formation on a substrate are considered in Section 4. Laser plume–ambient gas interaction, the conditions for the formation of different nanostructures, and mechanisms for cluster formation are the topics of Section 5. In Section 6, the formation of cluster-assembled films and cluster networks is considered. Finally, we discuss the perspectives of laser ablation for the controlled production of nanostructures and the possibilities for scaling this method for industrial applications.

2. CLUSTER FORMATION BY LASER ABLATION: EXPERIMENTAL INSTALLATIONS AND DIAGNOSTIC METHODS

In this section, we describe the experimental installations for the formation of different nanoclusters in which a laser is used for ablation of a target to produce a flow of hot atoms. The general principles for the design of these installations are the same. The laser hits a target, ablates it, and produces a hot atomic flow in an expanding plume. This flow then interacts with atoms of noble gas in the chamber where the nanoclusters are formed. Depending on the laser parameters and on the kind of nanocluster to be formed, additional

components, such as a furnace, diagnostic tools, etc., complement the major parts of the installation.

First, we briefly describe the lasers available for material ablation and nanocluster formation. Then we present the typical experimental installations used for different nanocluster formation. Finally, the experimental techniques for the *in-situ* diagnostics of cluster formation conditions and for the *ex-situ* characterization of the final product are described.

2.1. Lasers

For ablation of materials, a variety of lasers is now available, with wavelengths ranging from 0.15 to 10 μm and with pulse durations spanning from several tens of femtoseconds up to continuous wave mode operation (cw lasers). The pulse repetition rate ranges from several Hertz to 100 MHz. With these lasers, the evaporation and atomization of any material is possible. Commercially available solid state Nd:YAG (fundamental wavelength, $1\omega = 1064$ nm; second harmonic, $2\omega = 532$ nm), Ti:sapphire lasers ($1\omega = 800$ nm; $2\omega = 400$ nm), along with XeF (353 nm), XeCl (308 nm), KrF (248 nm), ArF (193 nm), and F_2 (153 nm) excimer gas lasers with pulse durations of 10–30 ns, energy per pulse 0.5–0.3 J, and repetition rates of 10–30 Hz, are conventionally used in many experimental installations [15, 18, 26]. Recently subpicosecond lasers (pulse duration, 50 fs to tens of picoseconds) with high repetition rates of 10 kHz–100 MHz have been demonstrated to be very efficient for target ablation and nanocluster production [31–33].

The "state of the art" powerful Nd:YVO laser, newly developed at the Australian National University, can operate over a wide range of wavelengths, intensities, and repetition rates in the megaHertz range. This laser provides a flexible tool for producing an ablated flow of hot atoms of any material with a controlled temperature and ablation rate. It can be used for high repetition rate laser ablation and for the formation and deposition of different nanostructures. The laser generates 12-ps pulses with variable repetition rates of 1.5, 2.6, 4.1, and 75 MHz in the first (1.064 μm) and second (532 nm) harmonics at an average power of 50 W [38]. Further updating to generate in the third (355 nm) and forth (266 nm) harmonics is in progress. A variable repetition rate was achieved by extension of the laser cavity using a multipass cell to increase the cavity length up to 100 m. The time gap between consecutive pulses is from 15 to 750 ns. The laser produces a continuous flow of hot atoms with an ablation rate of $\sim 10^{20}$ atoms/s. This means that the mixture of the ablation plume and the ambient gas (usually argon) cools down slowly between the pulses, and the temperature of the mixture oscillates around the equilibrium level during the formation process [39].

Therefore, the temperature and atomic density in the nanocluster formation zone are controlled entirely by the tunable laser parameters: pulse duration, energy per pulse, and repetition rate. The principal scheme of this laser is presented in Figure 1.

Several important parameters, such as pulse duration, wavelength, pulse repetition rate, and laser fluence (J/cm^2) and intensity (W/cm^2), should be defined in order to choose

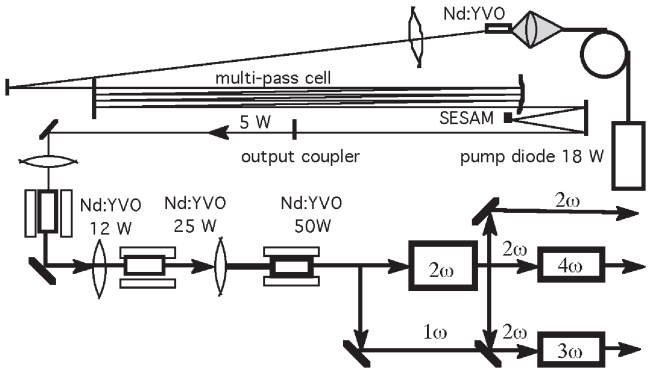


Figure 1. Principal scheme of a new powerful Nd:YVO laser installation at the Australian National University. The laser generates 12-ps pulses with variable repetition rates of 1.5, 2.6, and 4.1 MHz in the first (1.064 μm), second (532 nm), third (355 nm), and forth (266 nm) harmonics. It can be used for high repetition rate laser ablation and deposition of any material for the formation of different nanostructures under conditions controlled by flexible laser parameters.

an optimum laser for the ablation of a particular target and for cluster formation. First, the absorbed energy density delivered to the target should exceed the ablation threshold, which is in the range of 1–2 J/cm^2 for the majority of materials, and it can be defined for any material, as explained below in Section 3 [31]. The laser intensity and wavelength should be chosen appropriately to ensure efficient absorption of the laser radiation. Second, the molecular content and ionization state of the plume depend on the laser intensity distribution in time during the laser pulse and in space across the focal spot on the target surface.

2.2. Experimental Installations

2.2.1. Installation for the Production of Carbon Nanotubes

The use of laser ablation for nanocluster production was proposed by R. Smalley in 1983 [18]. The first installation used a commercial solid state low repetition rate laser with a noble gas flow through a tube placed in a furnace ($\sim$1200 K) where single-walled carbon nanotubes were produced. Later Smalley and co-workers [27] proposed the use of two lasers for enhancing the mass production of carbon nanotubes. Their current installation [40] uses two Nd:YAG lasers for ablation of a graphite target containing the catalysts cobalt and nickel (1 at.% each). The target is maintained at 1473 K in a tube furnace in an argon atmosphere (100-sccm flow, 66.7-kPa pressure, 500 Torr). The lasers produce 8-ns, 300-mJ pulses, with a repetition rate of 10 Hz. The first laser operates at the fundamental wavelength of 1064 nm, while the second laser operates at the second harmonic of 532 nm and hits the target 50 ns after the first laser shot. The laser beams are adjusted to travel colinearly such that they overlap on the target. The weblike deposit contains the nanoclusters that were produced in the laser plume and in argon gas. The deposit is collected either on the brass water-cooled collector that is positioned at the end of a quartz tube just outside the furnace or on the

quartz tube walls. The scheme of the installation is shown in Figure 2.

2.2.2. Experimental Setup for Formation of Metal and Covalent Clusters

The second typical installation is the so-called "laser vaporization cluster source" [1–3] that is used for production of metal and covalent clusters. In these devices, the clusters are also produced by the interaction of an ablated plume with the flow of a noble gas in the interaction chamber. The cluster flow is then released through a nozzle into a channel where mass (size) selection of clusters is performed. The flow of size-selected clusters is then deposited on a substrate. The scheme of a typical laser vaporization cluster source is shown in Figure 3.

2.2.3. Experimental Setup with High Repetition Rate Short-Pulse or Continuous-Wave Lasers: Formation of Cluster-Assembled Foams and Carbon and BN Nanotubes

The third type of installation is based upon commercially available powerful short-pulse high repetition rate lasers (see Fig. 4) [32, 33]. Short time intervals of 10–100 ns between the laser pulses lead to the overlapping of plumes produced by consecutive pulses. Proper tuning of both the laser intensity on the target surface and the pulse repetition rate allows one to control the temperature of the laser plume and noble gas mixture in the nanocluster formation zone, and a furnace for temperature control is not required. In installations described below, the laser parameters and scanning speed of the laser beam over the target surface are chosen to optimize the ablation rate, the temperature, and the molecular and ion states in the ablated plume.

Two high average power, high repetition rate Nd:YAG lasers ($\lambda = 1.064\ \mu$m) were used for laser ablation of a carbon target and for deposition experiments [32]. One laser

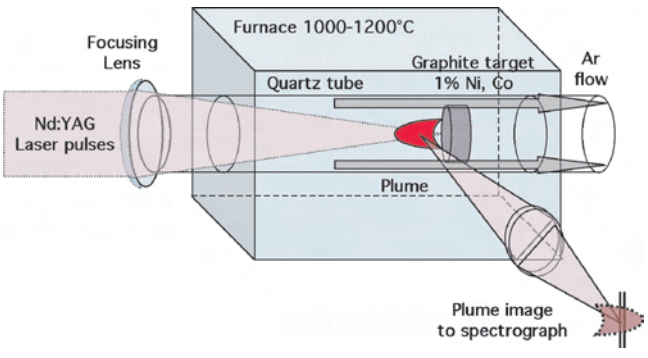


Figure 2. The first installation of single-walled carbon nanotube production proposed by R. E. Smalley [18]. The commercial solid state low repetition rate laser ablates a graphite target in a furnace heated to 1000–1200 °C with argon flowing through the furnace. Adapted with permission from [18], R. E. Smailey, *Laser Chem*. 2, 167 (1983). © 1983, ???.

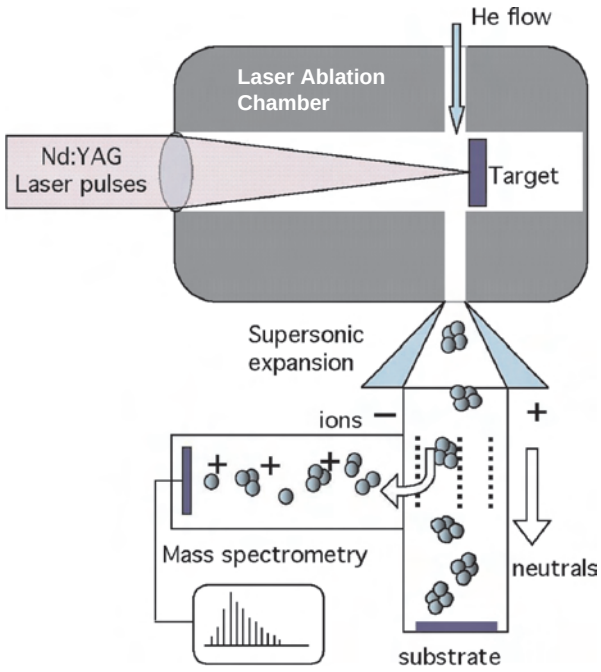


Figure 3. The scheme of a laser vaporization cluster source device that is used for the production of metal and covalent clusters, cluster-assembled film deposition, and mass spectrometry analysis. Adapted with permission from [1], K. Sattler, Ed., "Cluster Assembled Materials," Vol. 232 (1996). © 1996, Trans Tech Publications Limited.

produced 42 W in 120-ns-long pulses with a variable repetition rate in the range of 2–25 kHz. This laser created an average intensity of $\sim 2 \times 10^9$ W/cm^2 at the focal spot $S_{foc} = 1.6 \times 10^{-5}$ cm^2 on the target, or an energy density

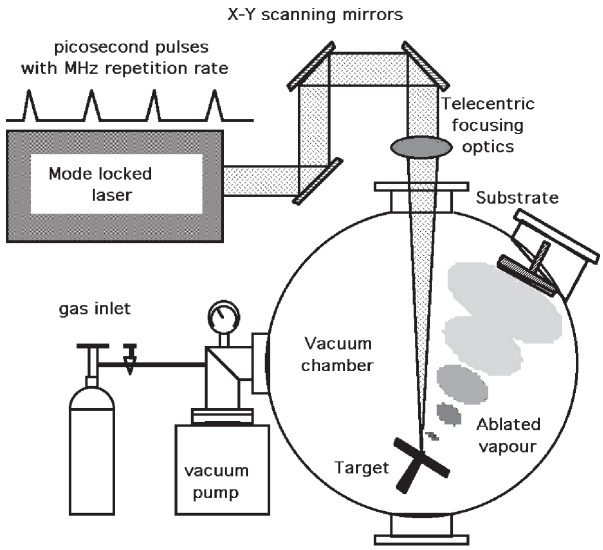


Figure 4. Schematic of the installation used for the formation of nanoclusters (carbon and boron nitride nanotubes) and nanoclustered materials (magnetic carbon nanofoam) in Ar at constant pressure with the use of high repetition rate short-pulse lasers. This installation is equipped with a powerful new laser [38] which allows the choice of all laser parameters: intensity, repetition rate, wavelength, etc., in order to obtain optimal conditions for the formation of the desired nanocluster.

in excess of 100 J/cm^2. The second one was a 25-W 60-ps mode-locked laser with a repetition rate of 76 MHz. Note that the period between the consequent pulses in this laser comprises only 13.2 ns. An intensity of $\sim 3 \times 10^9$ W/cm^2 was produced at the focal spot $S_{foc} = 1.8 \times 10^{-6}$ cm^2 with this laser [32].

2.3. *In-Situ* Characterization of Cluster Formation Conditions

2.3.1. Ablation Rate: The Feedstock for Cluster Formation

The rate of supply of hot atoms in the cluster formation zone depends on the laser ablation rate. This rate is conventionally measured by the target mass loss averaged over many laser pulses. This measurement provides an immediate check on how close the laser parameters are to those required for optimum evaporation. Indeed, the mass evaporated per single pulse can be estimated as the absorbed laser energy divided by the energy of evaporation per unit mass for a particular material. For example, if all the absorbed laser energy could be spent in evaporation, that is, spent only on breaking bonds, then 0.5 J of laser energy absorbed in a carbon target with the energy of vaporization at 30.4 kJ/g would result in ablation of 1.64×10^{-5} g of the target mass. However, the absorbed energy also goes into heat and shock waves in the target and in the kinetic and thermal energy of the ablated species.

Given optimum laser parameters, the ablation rate was found and measured in experiments with an Nd:YAG laser ($\lambda = 1.064$ μm; pulse duration, $t_p = 120$ ns; energy per pulse, $\sim$3 mJ at 10 kHz) [32]. The evaporation rate by Rode et al. [32] was deduced from the volume V_{evap} of the evaporated conelike crater on the target surface, measured with both an optical microscope and a scanning electron microscope for a given number of pulses. For a single laser pulse the measured rate comprises 5.6×10^{14} particles/pulse (or 0.01 μg/pulse). Note that the measured differential rate $R_{evap} = (1.4 \pm 0.2) \times 10^{26}$ atoms/(cm^2/s) qualitatively complies with the theoretical estimate of 10^{27} atoms/(cm^2/s) [31]. One can easily note the importance of the high repetition rate laser. For the 10-kHz repetition rate by Rode et al. [32], the ablation rate comprises 10^{-4} g/s. In this case, the carbon feedstock into the nanocluster formation zone amounts to 5.6×10^{18} particles/s.

A similar evaporation rate per pulse was observed in experiments where an XeCl excimer laser ($\lambda = 308$ nm; repetition rate, 10 Hz; $t_p = 17$ ns) was used for the ablation of graphite [41]. The ablation rate decreased after 100 laser pulses per single spot on the target surface.

A C–Ni–Co target was ablated by a single CO_2 laser pulse (20 ms; average intensity, 1.3×10^5 W/cm^2) in an argon atmosphere with the goal of carbon nanotube production [29]. The ablation rate was measured by the target weight loss of 0.1 mg after 10 laser shots, thus amounting to 10^{-5} g/pulse.

The influence of the scanning speed of a laser beam over the target surface on the evaporation rate also has to be considered. If the laser focal spot is kept at the same position on the target for several consecutive pulses, a crater is produced in the target, and its depth increases with each successive pulse. At very high repetition rates, this leads to an increase in the density of the vapor in the crater. As a result, the laser light is likely to be absorbed in the vapor away from the solid target surface and consequently, the interaction regime changes to a less favorable one in terms of evaporation efficiency [42]. When the crater depth becomes comparable with the diameter of the focal spot, the ablation rate drops in comparison to the single pulse regime. In a deep crater, the thermal losses from the vapor into the crater walls dominate the laser–target interaction process.

2.3.2. Emission Spectra of the Plume and Ambient Gas Mixture

Transient Temperature The emission spectrum of the target surface and the laser plume consists of two easily distinguished parts: one is continuous radiation, and another is line radiation. The average temperature of the mixture of the ablated plume with ambient gas can be found by fitting the continuous radiation spectrum with Plank's radiation law. Line radiation allows one to identify the molecular and ion content in the plume and to estimate the temperature and the level of ionization.

For example, time-integrated, space-resolved, target surface and plume temperatures were measured by Rode et al. [32] for a carbon target ablated by 10-kHz 120-ns laser pulses in a vacuum. The continuous radiation in the 600- to 800-nm range fits well with the Planck spectrum, with the temperature in the range of 2500 to 3500 K.

Single-wall carbon nanotubes were grown by ablation of a carbon–cobalt–nickel target with a Nd:YAG laser (1.064 μm, 300 mJ/pulse, 8 ns pulse width) in argon gas in an oven heated to 760–1100 °C [43]. The transient temperature of the carbon plume diffusing through the argon was estimated by Puretzky et al. [43] by fitting the measured blackbody radiation spectrum to Planck's law. The temperature of the carbon–argon gas mixture changes from 3500 °C at 0.1 ms to 1488 °C at 1 ms after ablation. It decreases to the oven temperature during the next 4 ms.

A similar procedure of fitting the continuous emission spectra to the Planck function by Kokai et al. [29] gives the temperatures for a carbon cluster–argon mixture in the range 3410–3540 K. Note that in this case, the single-wall carbon nanotubes were synthesized by ablation of a C–Ni–Co target by a 20-ms CO_2 laser single pulse ($\lambda = 10.6$ μm; peak power, $\sim$1 kW; average intensity, 1.3×10^5 W/cm^2) in an argon atmosphere. Thus, under all seemingly different conditions, the temperature in the formation zone for carbon nanotube formation is around $\sim$3000 K.

Molecular Content of a Plume Determined by Spectroscopic Measurements The carbon target was ablated by relatively long 120-ns laser pulses at a 10-kHz repetition rate at an intensity of $1–3 \times 10^9$ W/cm^2 [32]. The analysis of the line radiation spectrum of a carbon plume in a spectral range below 400 nm by Rode et al. [32] revealed that there is rather weak radiation from neutral carbon at 248 nm (the first excitation line for carbon) along with the intense lines corresponding to singly ionized and doubly ionized carbon. The temperature estimates led to high values of 1.5–6 eV in the plume.

The emission spectrum of a carbon–argon mixture was recorded by Arepalli et al. [40] in the spectral range of 300–650 nm. The spectral features were identified as emission of a Swan system from C_2 and a Comet Head system relevant to the presence of C_3. Comparison of the computed spectra to the measured excitation spectra allowed estimation of the vibration temperature in the range of 2500 to 4000 K. The temperature measurement results clearly demonstrate that the ionization time in the plume should be considered along with the intensity of the laser radiation on the target surface for predicting the plume temperature in the cluster formation zone [44].

2.4. *In-Situ* Characterization of the Cluster Size (Mass)

2.4.1. Cluster Mass Abundance Spectra

Clusters are formed as a result of interaction of the laser plume with noble gas in a cluster nucleation chamber (Fig. 3). The clusters then flow through a nozzle to the channel leading to the deposition chamber, where a substrate is located [2, 3, 45].

Cluster size (mass) distribution is conventionally determined by a time-of-flight method in a double-field configuration (see, for example [1–3] and references therein). In this method, the ionized clusters are deflected in the applied electric field and then separated in accordance with the charge/mass ratio. When neutral clusters are produced, a separate short-wave laser is used to ionize the neutral clusters, and then they are deviated by an applied electric field and analyzed with a spectrometer. The photon energy of the short-wave laser is chosen to be slightly larger than the first ionization potential of cluster material to assure that only single charged clusters will be produced. The mass abundance spectra have characteristic maxima at "magic numbers" which correspond to the most stable remarkable atomic structures, such as fullerene (C_{60}), cobalt and nickel icosahedra, aluminium octahedra, or metal clusters with closed electronic shells.

The most spectacular mass distribution (or, distribution of number of atoms per cluster) of carbon clusters produced by laser ablation of graphite was obtained by Rohlfing et al. [45] (see Fig. 5). The distribution is bimodal and the highest cluster intensity in the second peak corresponds to the famous fullerene, C_{60}, cluster. A typical mass resolution in this method, $M/\Delta M$, commonly stands between 100 and 1000, but it can be improved up to 30,000 [1].

2.4.2. In-Situ Absorption Spectroscopy of Carbon Nanoparticles: Average Cluster Size

In order to estimate the size of carbon clusters in the plume–ambient gas mixture absorption spectroscopy was used by Puretzky et al. [46]. The short (35–77 nm) single-wall carbon nanotubes were formed using nanosecond Nd:YAG laser ablation of a C–Ni–Co target inside a high-temperature (760–1100 °C) argon-filled oven. The light from a pulsed Xe lamp passed through the ablated plume in the oven, and then the extinction spectrum was measured. The set of extinction spectra was also calculated on the basis of Mie theory for the scattering and absorption of light on spherical particles of a relatively large size (diameter $d > 0.05\lambda$ where λ is the wavelength of the scattered radiation) [47]. For such particles, the extinction spectra are sensitive to the particle size. Comparison of calculated and measured spectra has shown that good fit of the experiment to the theory is obtained for the cluster diameter of 80 nm. This means that 1 ms after the ablation, when the spectra were measured, the carbon atoms have already aggregated into large clusters.

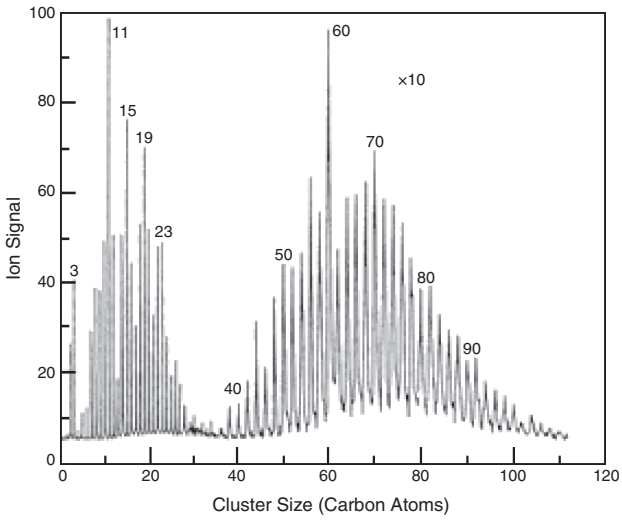


Figure 5. Mass distribution of carbon clusters, containing n carbons, created by the laser ablation of graphite. Reprinted with permission from [45], E. A. Rohlfing et al., *J. Chem. Phys.* 81, 3322 (1984). © 1984, American Institute of Physics.

2.4.3. Imaging of a Plume Diffusion Through the Ambient Gas: Onset of and Time for Cluster Formation

Carbon nanotubes are conventionally formed through interaction of a Laser-ablated plume of a carbon–cobalt–nickel target with a noble gas, which fills the experimental chamber. However, it is still unclear how long it takes for nanotube formation, when the growth process starts and ends, and what the conditions (temperature and density) in the formation zone are. Recently several experimental groups published results clarifying some of the stages in the cluster building process.

Single-wall carbon nanotubes were synthesized by ablation of a C–Ni–Co target with 20-ms pulses from a CO_2 laser (1 kW peak power; average intensity, 1.3×10^5 W/cm^2) in an argon atmosphere [29]. A high-speed video imaging technique was applied to record the expansion of the ablating plume. The authors observed that the moment when the carbonaceous material in the expanding plume became visible was 3 ms after the beginning of the laser pulse. This moment can be qualitatively identified as the moment when the number of big clusters became sufficient to make the absorption in the plume visible. The expansion velocity of

the plume in the argon atmosphere (the velocity of propagation of the diffusion front) was measured to be $\sim 10^3$ cm/s as estimated from the plume images at 1.7 and 3.3 ms.

Puretzky et al. [43] suggested that the time of nanotube growth is related to the time dwelt by the plume in the region of the furnace where the temperature is almost uniform. In order to measure this time dwell, the propagation of the laser-ablated C–Co–Ni plume diffusing through an Ar atmosphere at 500 Torr was registered *in-situ* by measuring the scattered light from a time-delayed probe laser beam (XeCl laser, 308 nm, 30 ns, 6 mJ/cm^2) [43]. By measuring the velocity of plume diffusion, the above dwelling time or formation time can be estimated. This time ranges from 15 to 120 ms. The authors [43] also measured the length distribution among nanotubes created under slightly different temperature conditions. The maximum length distribution of the nanotubes formed ranges from 35 to 77 nm. One can relate the above growth time to the corresponding nanotube length and obtain the growth rate for the nanotubes, which ranges from 0.6 to 5.1 μm/s. The growth rate measurements are in agreement with simple kinetic estimates of the single-walled nanotube formation rate under different thermal conditions and similar Ar pressure, which are in the range of 0.5–100 μm/s [48].

It is interesting to note that the measured velocity of the leading (sharp) edge of the diffusing plume in [43] changes abruptly at $\sim$3 ms, decreasing from $\nu_D = 8$ to 0.76×10^3 cm/s. This velocity is directly related to the average diffusion velocity, $D = l\nu/3$, of carbon atoms in the argon atmosphere (here $l = (n\sigma^{-1})$ is the cluster mean free path; n is the number density of particles in the formation zone (mainly argon atoms); $\sigma \sim r_{\text{cluster}}^2 \sim N^{2/3}$ is the cross section for cluster–argon collision, where r_{cluster} is the cluster radius and N is the number of atoms in the cluster). The velocity of a cluster decreases as the cluster mass increases since $\nu \sim (T/Nm)^{1/2} \sim N^{-1/2}$. Therefore, the diffusion velocity depends on the cluster size according to $D \sim N^{-7/6}$ [48]. The distance traveled by the diffusion front as it was measured in the experiment can be expressed as $R = (Dt)^{1/2} = \nu_{\text{D}} t$. Therefore $D \sim (\nu_{\text{D}})^2 t$. The cluster size grows as the diffusing front propagates through the cluster formation zone. Thus, in accordance with the plume propagation measurements in [43] after 3 ms, the number of particles in an average cluster increases $\sim$100-fold.

2.5. *Ex-Situ* Characterization of Individual Clusters and Cluster-Assembled Films (Foams): Structural and Electronic Properties

In this section we describe the diagnostic methods for *ex-situ* characterization of structural and electronic properties of individual nanoclusters and cluster-assembled materials.

2.5.1. Electron Microscopy

A very powerful tool for studying the electronic and structural features of nanoclusters is electron microscopy, with which imaging of the electron beam scattering on and transmission through the studied structure as well as the interference and diffraction patterns can be investigated. Electron microscopy allows the identification of subtle structural features of complicated individual nanoclusters and nanocluster networks. The images of single-wall carbon nanotubes were obtained by using scanning tunneling microscopy [49, 50]. This technique gives atomically resolved lattice images of relatively high contrast. However, there are difficulties in observing such delicate features as the chirality of nanotubes by using this technique. Recently, high-resolution transmission electron microscopy, operating in the imaging mode, was used for characterization of single-walled carbon nanotubes and nanotube bundles produced by the laser ablation technique [51]. A high-resolution transmission analytical electron microscope, the JEM-300F (JEOL) equipped with a 300-keV field emission gun (NIRIM Tsukuba, Japan), was used for these observations. The microscope has a theoretical spatial resolution limit of 0.17 nm. In reality, lattice fringes of 0.21 nm were resolved in the "core" region of a single-walled nanotube, as is seen in Figure 6. Considering the orientation and contrast of these fringes, a researcher is able to observe and estimate such subtle nanotube characteristics as chirality (or helicity) [51].

2.5.2. Raman Spectroscopy

The Raman scattering technique is a powerful tool for determining the structural features of different crystalline structures, nanoclusters, and nanomaterials. Comparison of the known Raman spectra of a bulk crystalline material to that of the nanoclustered one allows the determination of a transition to a nanostate where quantum confinement effects can be observed. The Raman spectra of carbon nanoclustered materials are usually compared to crystalline diamond and graphite lines. Diamond has a very sharp peak at 1330 cm^{-1}, and graphite is characterized by two broad peaks at 1400 cm^{-1} (D-peak) and 1600 cm^{-1} (G-peak) [52]. Amorphous carbon (a-C), regardless of the way it is produced, tends to have a broad single asymmetric peak with a maximum around 1500–1550 cm^{-1} [52, 53].

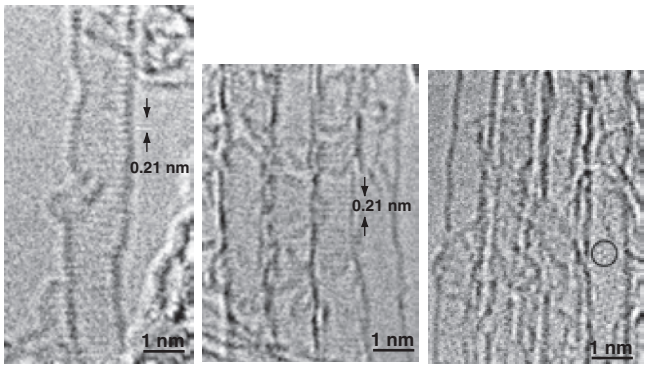


Figure 6. High-resolution transmission electron microscope (HRTEM) images of an individual single-wall carbon nanotube isolated (left) and in a bundle (center) exhibiting 0.21-nm fringes in the core region in agreement with a "zigzag" arrangement of atoms. A 2D dot contrast in the core region, in agreement with the armchair arrangement of the atoms in a tubular shell, is shown in the right image. Reprinted with permission from [51], D. Goldberg et al., *Carbon* 37, 1858 (1999). © 1999, Elsevier Science.

Raman spectra of the carbon nanofoam measured with an excitation wavelength of 632.8 nm showed two separated broad peaks centered at 1341–1353 and at 1569–1579 cm^{-1}, which were attributed to graphitic sp^2 bonding [54]. The detected peaks at 1350 and 1580 cm^{-1} were assigned to disordered and ordered carbon phases, respectively, and the samples were described as glassy carbon with very little order.

A resonant Raman scattering process on single-wall carbon nanotubes was observed with the laser excitation wavelength varying from 514.5 to 1320 nm [55]. The resonance results from the one-dimensional quantum confinement of the electrons in the nanotube. It was shown theoretically and later confirmed experimentally [56] that depending on the symmetry and diameter of the carbon nanotubes, they can be either metallic or semiconducting. These quantum size effects manifest themselves in the Raman scattering signal from the nanotube vibrational modes. In [55] carbonaceous structures were produced by laser ablation. These nanotubes with a very narrow diameter distribution were separated through a long purification procedure, including a microfiltration technique to remove most of the carbon nanoparticles which were also produced in the process.

According to theory [56], nanotubes are described by two indices (n, m). These indices (n, m) define the atomic coordinates for the one-dimensional unit cell of the nanotube. The tube with $n \neq m \neq 0$ has a chiral symmetry. The achiral tubes ($m = 0$, or $m = n$) are divided into two subclasses: tubes with $(n, 0)$ are referred to as "zigzag" tubes and tubes with (n, n) are referred to as "armchair" tubes. The Raman data were analyzed assuming only the armchair symmetry for the studied tubes because the previous calculations have shown that chiral tubes exhibit very few Raman-active modes in the range 400–1500 cm^{-1}.

The theory predicts a strong diameter-dependent Raman line: 206 cm^{-1} for (8, 8) nanotubes with diameter $d =$ 10.85 Å and 150 cm^{-1} for (11, 11) nanotubes with diameter $d = 14.927$ Å. Using 514.5-nm laser excitation, a strong broad line was observed at 186 cm^{-1}. This line associated with a randomly oriented mixture of nanotubes with diameters corresponding to (n, n) tubes with n ranging from 8 to 11. Thus, the Raman observation allows a reasonable estimation of the average diameter of nanotubes in the sample.

It was also observed [55] that the Raman spectrum (intensity and positions of the peaks) strongly depends on the laser excitation frequency. The redistribution of the spectral intensity as a function of the excitation wavelength is an indication of unusual resonant Raman scattering. Resonant scattering occurs when the energy of the incident photon matches the energy difference in an electronic transition due to optical absorption. It is now well known that starting from some small dimension on the nanometer scale (a certain diameter of nanotube or nanocluster), the energy difference between the electronic states responsible for the optical absorption increases when the nanocluster diameter decreases. This effect is often referred to as the quantum size effect or as a one-dimensional quantum confinement phenomenon. For a large diameter, the character of the electronic states is essentially independent of the tube diameter, so the Raman spectrum should be very close to that for a flat graphene sheet and should be independent of the excitation frequency. The experimental observations and calculations of a Raman spectrum presented by Rao et al. [55] provide strong evidence of a diameter-dependent Raman effect that is a consequence of quantum confinement in single-wall carbon nanotubes with diameters of 10–15 A.

2.5.3. Electron Energy Loss Spectroscopy

The character of the bonding in an individual nanocluster or in a nanocluster-assembled material can be investigated using the electron energy-loss spectra (EELS) technique [57]. For example, in carbonaceous materials the important characteristic of the electronic structure is the ratio of sp^2 hybridized bonds (signature of graphite) to sp^3 bonds (signature of diamond). This ratio $sp^2/(sp^2 + sp^3)$ in carbon nanofoam was investigated by the EELS technique [37]. The EELS spectra were analyzed using a GATAN 666 parallel electron energy-loss spectrometer (PEELS). PEELS spectra were collected by two methods. First, a number of different regions of laser-produced carbon nanoclusters were chosen and five different spectra were collected from around each region. In the second method, in-house software was used to collect a number of spectra along a chosen straight line across the nanocluster. This method allowed one mapping of the variation of sp^2 bonding across a single cluster with 1-nm spatial resolution [58]. The mapping of a 6-nm carbon cluster, which is a typical structural unit of the nanofoam, revealed that the carbon atoms inside the clusters have predominantly sp^2 bonding, while at the cluster boundaries sp^3 "diamond-like" bonding prevails.

2.5.4. Surface Area Measurements for Nanocluster-Assembled Materials

Nanocluster-assembled materials usually have a porous, sponge-like or foam-like structure. One of the characteristics for such materials is the characteristic surface area, that is, the area of the inner pores for the unit weight of material. The multipoint Brunauer–Emmett–Teller (BET) method [59] for surface area measurements is commonly used to measure this parameter.

Using this method, the surface area measurements for carbon nanofoam produced by the ultrafast laser ablation of graphite were performed using a surface-area analyzer [37]. About 100 mg of nanomaterial was placed into a sample holder and first outgassed at 120 °C under a nitrogen flow and immersed in liquid nitrogen. The measurements showed that the surface area of the nanofoam is in the range of 300–400 m^2/g. Detailed studies of the desorption of the nitrogen from the foam can be made using the BJH method [60], which calculates the pore volume distribution. The analysis of nanofoams formed at different argon pressures demonstrated that all foam samples have essentially constant pore volume, with no evidence of a preferred pore size.

2.5.5. Small-Angle X-Ray Scattering

Small-angle X-ray scattering data have been found useful for analysis of the fractal dimensions of aerogels, carbon nanofoam [37], and nanoclusters [1–3]. X-rays from a conventional Cu X-ray source were collimated with standard

Guinier optics onto a carbon nanofoam sample to reveal its structural features. The scattered radiation has been collected on a home-built camera [37]. According to Schmidt [61], fractal structures of fractal dimension d give rise to a power law decay of the scattered intensity $I(q) \sim q^{-d}$, where q is the scattering vector. The exact location of this power law regime on the q scale is in practice delicate. However, it is possible in some cases to identify this functional form over a reasonable q range. Rode et al. [37] found the power law to be valid in a range of 0.1–0.4 Å^{-1}, leading to an estimation of the fractal dimension of 2.4. This means that the atomic arrangement in space lies somewhere in between homogeneous two-dimensional sheets and three-dimensional (3D) volumes.

2.5.6. Measurement of the Bulk Density of Cluster-Assembled Materials

A convenient method for determining of the bulk density of cluster-assembled materials (foams, aerogels, sponges, etc.) is Rutherford backscattering [62]. The bulk density of the carbon nanofoam deposited on a silicon substrate [37] was determined from 2-MeV He^+ ion backscattering measurements. The area density of the sample was deduced from the energy width of the carbon peak and the energy shift of the silicon edge, and the thickness of the foam sample was obtained from the scanning electron microscope images (see Fig. 7), giving the density of the foam as $2–10 \times 10^{-3}$ g/cm^3.

3. HOT VAPOR (PLASMA) PRODUCED BY LASER ABLATION

In this section we underline the main physical features of laser–matter interaction and formation of an evaporated plume of hot vapor and plasma. The simple and self-consistent procedure allows the choice of the proper combination of laser and material parameters to design an experiment for producing a plume with the desired parameters, for example, to produce an atomic flow by laser ablation.

3.1. Laser–Matter Interaction: Light Absorption and Ablation Threshold

In order to remove an atom from a solid by means of a laser pulse one should deliver energy in excess of the binding energy ε_b of that atom (the energy of evaporation per atom).

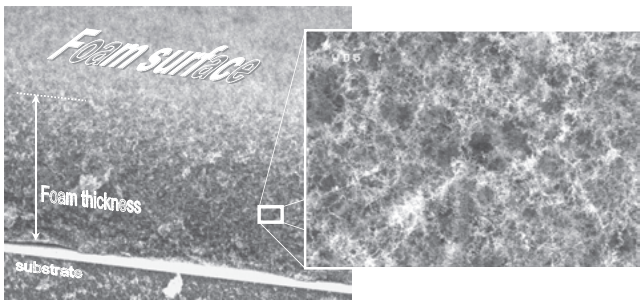


Figure 7. Scanning electron microscope images of a 100-μm-thick nanofoam assembled from carbon nanoclusters and deposited on a silicon substrate.

Therefore, to begin with, efficient absorption of the laser energy is necessary. The ratio of the absorbed energy flux to the incident laser flux, or the absorption coefficient, depends on the material properties and on the intensity and wavelength of the laser.

Absorption coefficients for different materials and for various wavelengths can be found in the literature [31, 63, 64]. To illustrate spectral dependencies we present simple estimates for the absorption coefficient for metals. The laser field penetrates into a solid at the skin depth, l_s, which in metals is much smaller than the wavelength of the laser, $l_s \ll \lambda$. The skin depth, $l_s = \frac{c}{\omega k}$, is conventionally expressed through the imaginary part of the refractive index, k, and the laser frequency, ω [63]. One can show that the absorption coefficient is given by $A \approx \frac{4\pi l_s}{\lambda}$ [65]. As can be seen, the use of short wavelength lasers is preferable for good absorption. Electrons absorb the laser energy at a distance comparable to the skin depth. Suppose that the laser intensity distribution, I_0 (W/cm^2), over the focal spot diameter and during the pulse time, t_p, has a "tophat" profile. The absorbed laser energy per unit area F (J/cm^2) is expressed as $F = AI_0t_p$. This quantity is called the absorbed fluence.

The absorbed energy is confined in the electrons, and therefore, the maximum electron temperature at the end of the laser pulse is proportional to the absorbed laser fluence as

$$T_e = \frac{4F}{3n_e l_s}$$

where n_e is the number density of free electrons in the metal. Finally the hot electrons transfer their energy to cold ions, and the electron and ion temperature T_i equilibrates, such that $T_e = T_i$. The ions can leave the solid (or to be ablated) if $T_i > \varepsilon_b$. Now one can immediately obtain a relation between the laser fluence and material parameters needed to reach the ablation threshold $F_{threshold} = \frac{3}{4} n_e l_s \varepsilon_b$. A typical binding energy for the majority of solids is a few electron volts and therefore, the typical absorbed laser fluence for ablation threshold is around 0.5 J/cm^2.

One can see that ablation by long or short laser pulses occurs under very different conditions. In order to ablate the same amount of material with a short pulse, one should apply a larger laser intensity approximately in inverse proportion to the pulse duration. For example, laser ablation with a 100-fs pulse requires the intensity to be above 10^{13} W/cm^2 [31–33, 65], while with 10-ns pulses, the same material is ablated at much lower intensities, $\sim 10^8$–10^9 W/cm^2 [31]. The depth of material ablated per single short laser pulse is proportional to the skin depth, whereas in the case of long pulse ablation, the characteristic ablation depth is proportional to the heat conduction length per pulse $\sim (at_p)^{1/2}$, where a is the thermal diffusivity (cm^2/s) of the target material. Correspondingly, the number of particles evaporated per pulse differs by several orders of magnitude.

It has been shown [33, 65] that at laser intensities in excess of 10^{13} W/cm^2, practically any target material is ionized during the subpicosecond pulse time. Following ionization, the laser energy is efficiently absorbed by free electrons due to

inverse Brehmsstrahlung and resonance absorption mechanisms and does not depend on the initial state of the target. Consequently, the interaction of a laser pulse with both metals and dielectrics proceeds in a similar way. The efficient evaporation of dielectrics and the formation of atomic beams of ablated atoms is now possible using short and intense laser pulses. Since additional energy, ~10 eV/atom, is required for ionization, it is obvious that the ablation threshold for dielectrics must be higher than that for metals. Indeed, the experiments show that the ablation threshold for dielectrics is typically ~2 J/cm^2 [33].

3.2. Ablation Rate and Deposition Rate

The most efficient ablation will be achieved if all the absorbed energy is used in ablation only, and any losses and overheating of the plume are avoided. In this case, the instantaneous ablation rate (the ablation rate during the pulse) can be easily estimated as the absorbed laser intensity divided by the binding energy

$$(n\nu)_{\text{abl}} \approx \frac{AI_0}{\varepsilon_{\text{b}}} \ (\text{atoms/cm}^2/\text{s})$$

In fact this ablation rate is close to the experimentally measured value for the optimum ablation mode. This is the highest rate from all available ablation methods. The values for the instantaneous ablation rate span from 10^{32} atoms/cm^2/s for short, ~100-fs pulses, to 10^{27} atoms/cm^2/s for the nanosecond pulses. The total number of atoms evaporated per single pulse with the total energy E_{las} is given by $AE_{\text{las}}/\varepsilon_{\text{b}}$. The laser energy per short pulse is in the order of ~μJ, while in the nanosecond pulses it reaches 1 J. Thus, the ablation rate per pulse, with the pulse duration ranging from nanoseconds to femtoseconds, is correspondingly ~10^{12}–10^{18} atoms/pulse. Taking into account the different pulse repetition rates (up to 100 MHz for the short pulses and 10–30 Hz for the long nanosecond pulses) one can see that the average flow of laser-ablated atoms is $N \sim 10^{19}$–10^{20} atoms/s near the ablation surface. This figure, along with the temperature and ionization state of the vapor, defines the initial conditions for the formation of new structures in a chamber.

Let us consider the use of the atomic flow for the deposition of thin films on a substrate placed at a distance d from the target in an evacuated chamber. The film growth rate for the formation of a film with a number density n_{a} is defined as $N/2\pi d^2 n_{\text{a}}$. Thus, taking the average target to substrate distance of 10 cm and the number density of solid 10^{23} atoms/cm^3 the deposition rate for the formation of amorphous films with laser ablation can reach several tens of nanometers per second.

3.3. Criterion for the Full Atomization of Ablated Vapor

The phase state of the ablated vapors is determined by the conditions of evaporation, that is, it depends on the amount of laser energy absorbed in the ablating material. The energy delivered by a laser is spent on the breaking of interatomic bonds and on the kinetic energy of atomic expansion. If this kinetic energy is low, the vapor expands slowly, and therefore it can be condensed back near the target shortly after the laser pulse termination. We shall determine the magnitude of the kinetic energy sufficient for keeping the plume expanding into a vacuum in a gas state of non-interacting atoms. One can estimate the energy threshold necessary to achieve full atomization of the ablated plume on the basis of thermodynamic arguments similar to those used for establishing the criterion for complete vaporization of a material in an unloading stage after the action of a strong shock wave [66]. In this case all the nonequilibrium processes related to the energy transfer to ions are assumed to be completed before the expansion begins. The target material after the termination of the pulse experiences adiabatic expansion.

The adiabatic expansion from a solid state to the gas state can be described by the conventional relation between the gas energy, pressure P, and gas volume, V, but with a volume-dependent exponent (so called the Gruneisen coefficient). In order to get the final stage of expansion as a gas state, the adiabatic curve in the PV plane must pass higher than the critical point, $P_{\text{cr}}, V_{\text{cr}}$, separating the states of a homogeneous phase (atomized vapor) from states with a mixture of phases (gas and condensed liquid droplets). Given these constraints, it can be determined that the energy delivered by the laser per ablated atom must be approximately three to four times larger than the binding energy. Therefore, the absorbed laser energy density necessary to transform the ablated material into an atomized vapor should exceed the ablation threshold in accordance with the condition $F > 4F_{\text{thr}}$.

3.4. Damage and Condensation Thresholds

If the total energy delivered by the laser is close to the binding energy, $\varepsilon_{\text{total}} \sim \varepsilon_{\text{b}}$, then a solid experiences only a small density decrease from the standard solid density of the target material. The pressure in the material is comparable to the bulk modulus. Therefore, the final state of the target affected by the laser at this energy level might be considered "damaged," having cracks, flakes, etc. in the laser focal spot depending on the initial state of the target (such as the presence of defects, impurities, etc.). If the deposited energy is in the range $\varepsilon_{\text{b}} < \varepsilon_{\text{total}} < (3\text{–}4)\varepsilon_b$, then the final state of the expanding vapors may lie in the region of the pressure–volume parameter space where the mixture of phases is energetically favorable. The condensation of vapor into liquid droplets in a course of expansion in a vacuum may occur when the absorbed laser energy is within $F_{\text{thr}} < F < (3\text{–}4)F_{\text{thr}}$.

3.5. Ionization State and Velocity Spectrum of Ablated Ions

Ionization of a solid target material during the ablation process and the ionization of ablated vapors occurs due to processes of photoionization, multiphoton ionization, and ionization by electron impact (avalanche ionization) [66]. As ionization proceeds, the vapor is converted into a high-density plasma whose properties can eventually dominate the physics of the laser–vapor interaction. If the laser–matter interaction occurs at a high intensity, which is typical in ablation by femtosecond laser pulses, the full first ionization is completed during the first few femtoseconds at the

beginning of the laser pulse [65–68]. Afterward the interaction proceeds with the plasma. The energy distribution of the ions in the ablated flow appears to be non-Maxwellian. It can be roughly approximated as a two-bump distribution composed of the slow and fast parts. Experimental and theoretical studies have demonstrated that the energy distribution of outflowing ions is highly dependent on the ratio of the maximum intensity in the main laser pulse to the intensity in the prepulse (the so-called contrast ratio) [69]. It has been observed that for 100-fs laser pulses in the laser intensity range of 10^{14}–10^{16} W/cm^2, the ion energies vary from 100 eV to 20 keV [69]. Therefore, the interaction of an ionized flow, either with an ambient gas in the chamber resulting in cluster formation or with a substrate leading to the formation of a nanofilm, proceeds in a different way, as has been confirmed by experiments which are discussed below. We should stress that all parameters of a hot atomic flow which affect the process of a nanocluster formation can be controlled and steered in a desirable manner.

4. FORMATION OF NANOCLUSTERS IN A VACUUM

The whole process of nanostructure formation can be separated tentatively into the following stages. First, laser-ablated atomic vapors (or plasma) are produced to form a plume. Then, the laser plume expands into a vacuum and cools in the experimental chamber. The nanostructure formation process (nucleation) is composed of vapor–vapor, vapor–filling gas, vapor–substrate, and cluster–substrate interactions. Finally, the resulting product is annealed and cooled. In this section, we concentrate on the processes of nanostructure formation, taking the parameters of ablated vapors (plasma) as the initial conditions. Then, we discuss the characterization of the structure produced and relate the structural and material features to the formation conditions resulting from vapor–gas–substrate coupling.

We start from the most studied process of thin film deposition in an evacuated chamber. Recently a huge variety of different structures were deposited by the laser ablation method including refractory and complex materials. The literature covering this topic comprises thousands of references [i.e., 70, 71]. Our goal in this review is to point out relations between the properties of the atomic beam, the ambient gas, and the conditions created in the chamber to the properties of the emerging nanostructure. Such relations, when properly understood, would allow the control and prediction of the formation process. Therefore, we concentrate on several structures where this relation has been at least partly revealed. We also mention recent achievements in the quality and in the efficiency of thin film deposition, which makes the pulsed laser deposition technique more attractive for industrial applications.

4.1. Deposition of Nanometric Films in a Vacuum

Laser ablation has been successfully used for thin film deposition during the last 3 decades with the use of conventional nanosecond, 10- to 30-Hz lasers [70–72]. However, due to poor control over the laser beam parameters in these lasers, the surface quality (abundance of droplets) and material properties (for example, the sp^2/sp^3 bond ratio in carbon films) would differ from one laser installation to another with seemingly similar parameters. In addition, the deposition rate with lasers running at a repetition rate of 10–30 Hz was too low for most industrial applications.

The recent advent of short-pulse, high average power, high repetition rate lasers resulted in the elimination of these drawbacks of laser deposition. Progress in the understanding of the ablation process led to control over the pulse time shape (elimination of the prepulse), over the spatial distribution of the intensity across the laser focal spot (top-hat distribution), and over the temperature of the ablated ions. Now full control of the ablated plume parameters, at least in principle, can be provided. Future developments will determine whether it will be possible to make this process cost effective and to scale it to the industrial level.

4.1.1. Interaction of Hot Atoms with a Substrate in a Vacuum

Let us describe the succession of processes accompanying thin film deposition in a vacuum. We consider, for example, the flow of carbon atoms forming diamond-like carbon films on a substrate (conventionally the substrate is a silicon or silica) in a high vacuum, $P = \sim 10^{-6}$–10^{-7} Torr, which corresponds to a density of air molecules of $\sim 2 \times 10^9$–10^{10} cm^{-3}. A short-pulse laser with a high repetition rate produces 10^{19}–10^{20} atoms/s, which destroys the vacuum after several seconds of operation [32, 73–75]. Therefore, continuous chamber evacuation is necessary in order to maintain constant conditions during the long, 30- to 120-min deposition process. Continuous chamber evacuation at a rate of 2×10^3 liters/s maintains the number density of particles in the chamber at an approximately constant level of $n_a \sim 2 \times 10^{10}$ cm^{-3} [73]. Under these conditions, the mean free path for the particles (carbon atoms and air molecules) is $l \sim 1/n_a\sigma = 5 \times 10^4$ cm, assuming that the cross section of atom–atom collisions is $\sigma \sim 10^{-15}$ cm^2. Hence, one can suggest that there is no influence from any collisions in the chamber on the film formation process at the substrate and no cooling of the carbon atoms under this vacuum condition. Thus, the films of any material are formed due to direct vapor–substrate interaction at the vapor's temperature, which approximately corresponds to the target surface temperature. Note that these conditions are appropriate for the formation of sp^3 carbon bonds assuming $T > 10^3$ K. However, not all of the collisions between the ablated atoms and the substrate lead to the sticky attachment of an atom to the substrate. Some atoms can rebound from the substrate or from the chamber walls, cool down, and flow around the vacuum chamber [3, 73, 76]. Some small deposition of the ablated material on the rear side of the substrate may serve as direct experimental evidence of this effect.

Another important factor for deposited film formation is the pressure on the substrate during the film building process. This pressure consists of two components: kinetic pressure, P_{kin}, and thermal pressure, P_{th}. The momentum transferred to the substrate by the incoming flux of atoms (ions) determines the kinetic pressure. However, the kinetic

pressure is small in comparison to the thermal pressure, which depends on the average temperature of the substrate. The average temperature on the substrate surface generated by the impact of the ablated flow can be calculated through the maximum temperature on the target surface, T_{max}, the laser repetition rate, and the pulse duration [73]. The thermal pressure and average temperature on the substrate were estimated as 20 kbar and 1600 K, respectively, for the 120-ns pulse laser and 40 kbar and 3000 K for the 60-ps laser by Gamaly et al. (1999, 2000) [31, 73]. These pressure and temperature values are close to the range of parameters most appropriate for the carbon-to-diamond phase transition to occur under equilibrium conditions.

One can see that an increase in laser intensity will increase the energy of the ablated ions and consequently will increase the temperature and pressure on the surface of the substrate, thus affecting the conditions for thin film formation.

4.1.2. Experiments: The Material Properties of the Films Produced

High average power, high repetition rate Nd:YAG lasers ($\lambda = 1.064\ \mu$m) were used for laser ablation of carbonaceous targets and deposition experiments [73, 75]. Laser radiation with parameters of 60 ps, 3×10^9 W/cm^2, and 76 MHz is absorbed well ($\sim$85%) by a graphite target and produces an almost fully ionized laser plume with a surface temperature (average over many pulses) in the range from 2500 to 3500 K [73]. Thin carbon film deposition was performed in a chamber evacuated to 10^{-7} Torr on a mica substrate. The surface morphology of carbon films with a thickness of 25–120 nm has been investigated with an atomic force microscope [73–75, 77]. Films with thicknesses of 20–25 nm appear to be almost atomically smooth with the surface root mean square (rms) roughness around 0.4–1 nm [73, 75]. The optical bandgap of 0.5–0.7 eV extracted from ellipsometry measurements and Raman spectra indicated that the deposited films had properties similar to the diamond-like amorphous carbon (a-C) films.

Studies of the influence of laser intensity in the range 10^{14}–10^{16} W/cm^2 and wavelength in a range of 400–1000 nm on the properties of deposited films unveiled several interesting phenomena [73–76]. The energy distribution of ions which were ablated with short (around 100 fs), intense (in excess of 10^{14} W/cm^2), and prepulse free laser pulses has a pronounced two-bump form composed of a main thermal part, E_{th}, and a fast part, with energy 3–5 E_{th} [74]. Studies of the deposited films revealed that the transmission of the films increases with increasing laser intensity, while the ratio of sp^3 to sp^2 bonds decreases, being evidence of a less diamond-like character. The bandgap for these films determined with ellipsometry measurements was 0.85 eV, which supports the assumption of the amorphous nature of the films. Experiments using an intense short-pulse, high repetition rate laser demonstrated, along with atomic surface quality, an extremely high deposition rate up to a few micrometers/hour [73, 75], exceeding that for all other available methods of thin film deposition. It is instructive to compare with a deposition rate using a long pulse ($\sim$17 ns), low repetition rate (10–20 Hz) KrF laser (248 nm, 5–125 J/cm^2), where the deposition rate for the production of amorphous carbon films with high surface quality (rms roughness of 0.6 nm) constitutes 0.3–0.6 micrometers/hour [77].

4.2. Deposition of Complex Films

Laser ablation can be used to produce films of complex materials with the same stoichiometry as the original target. Provided that the laser intensity is chosen such that the bonds with the highest binding energy can be destroyed, laser ablation can produce an atomized beam containing the whole mix of constituent atoms. As a result, the laser plume contains a mixture of atoms with exactly the same atomic content and with the same atomic ratio as in an original target. Therefore the deposition process maintains the same stoichiometry in the deposited film or in a nanocluster as in the ablated target. As an example of complex structures being successfully deposited, we mention chalcogenide glasses (As_2S_3) [34] and high-temperature superconducting ceramics [78–80].

5. NANOCLUSTERS PRODUCED THROUGH THE INTERACTION OF ABLATED VAPOR WITH NOBLE GAS

The ideal mode of formation for any nanostructure is bottom-up or atom-to-atom attachment in a proper place in space, and at the proper rate in time, in order to form clusters with desirable characteristics. To approach this mode, one should be able to control the formation process on a space scale of angstroms and on a time scale less than a picosecond.

An atomic beam for cluster formation can be created by laser ablation with short (picosecond and subpicosecond) successive laser pulses, with the repetition rate up to 100 MHz. In this case, the shortest controllable time scales are the laser pulse duration and a time gap between the pulses. The smallest controllable number of atoms is the number of atoms evaporated per single laser pulse. Of course, the decrease of a single pulse duration and increase of the repetition rate will lead to a smaller controllable time scale and number of particles per pulse. In principle, there is no upper (lower) limit: in the ultimate limit of extremely short pulses and the highest possible repetition rate, one has atom-by-atom evaporation with as short as necessary time control and consequently an atom-to-atom attachment process of cluster formation.

The process of nanocluster formation is composed of several stages, namely, the formation of a laser-produced flow of hot atomic vapors, plume propagation, diffusion, cooling/heating in a chamber, a cluster nucleation process, and a final stage of annealing, cooling, and stabilization. In what follows, we describe all the formation stages of individual clusters and compare with experimental data where available. We briefly discuss size-dependent cluster material properties and the internal structure of nanoclusters. We then discuss interactions with the substrate and the formation of cluster-assembled films.

5.1. Nucleation, Growth, and Annealing

It is instructive to consider the simplest case of the formation of clusters composed from one element. The most studied cases are those of carbon cluster and metal cluster formation. We consider, for example, the thoroughly studied carbon cluster formation process in a noble gas filling of the experimental chamber. There are also many studies of cluster formation in a reactive gas atmosphere [7–9, 81]. The processes of cluster formation in reactive gases are more complicated and less developed. Therefore, we restrict ourselves to interactions between atomic beams with noble gas. Many features of the formation process that we describe are common for the formation of clusters of different elements.

We should note that many essential details of cluster formation mechanisms have still not been uncovered, and some of the proposed mechanisms are questionable. There is no clear understanding or mutual agreement about the onset of the nucleation process, the progression toward cluster building, or growth termination. Many theoretical models have been proposed. But due to the lack of *in-situ* diagnostics, no experimental validation of the proposed mechanisms has been obtained to date. However, the general picture of cluster formation can be presented in a reasonable scenario. In the following we present a qualitative picture of cluster formation based on kinetic and thermodynamic arguments.

5.1.1. Nucleation and Growth

The cluster formation mechanism appears to be simple from first appearances. The hot vapor of ablated atoms rapidly expands in a cold gas. Thus, the thermodynamic approach of fast vapor condensation into minidroplets, that is, clusters, seems straightforward. Close examination, however, reveals that the process is too fast in comparison to the equilibration time. The vapor density is relatively low, and the atom-to-atom collision time is not negligibly small in comparison to the characteristic time of the process. Thus caution should be exercised in applying the thermodynamic approach for describing the cluster formation process. In the following, we discuss and compare the thermodynamic and kinetic descriptions of the cluster formation process.

Thermodynamic Approach—Condensation Nucleation is a process through which clusters of a new phase grow inside an initial phase, which has been heated above the phase transition temperature. In thermodynamics the nucleation process describes the early stages in phase transformation, such as condensation of gases during expansion or solidification of melts. The expanding vapor (plasma) of ablated atoms during interaction with a cold gas or substrate can be transformed into a supersaturated state. This state is thermodynamically unstable. Therefore, a vapor–liquid phase transition develops leading to the formation of nuclei of the new phase, minidroplets of liquid. If some fast annealing process can suddenly terminate further growth of these droplets, then nanoclusters can be formed.

The nuclei of the new phase are formed due to fluctuations in an unstable system. A nucleus can continue its growth only in cases where its radius exceeds a definite critical value. Classical condensation theory [82] predicts well the critical size of a nucleus when a liquid droplet forms, starting the fast process of the phase transformation under the conditions of local thermodynamic equilibrium.

The small seeds (nuclei) of a new phase are created in the overheated, initially homogeneous, vapor phase due to fluctuations [82]. Such seeds are unstable structures because the formation of an interface between two phases is energetically unfavorable: to create a seed, the work to overcome the surface tension for the formation of a nucleus needs to be spent. The seeds with a size less than the critical value decay into the initial phase. The seeds with a size exceeding the minimum critical radius start to grow up, rapidly becoming the centers of the fast transformation into a new phase. The critical radius of a seed can be obtained from the condition of minimization of the free energy of a seed assuming that it has the form of a spherical droplet. Physically, this condition is equivalent to the balance between the surface tension and the thermal pressure. The critical radius of the seed is expressed through the temperature of the overheated phase, T, as [82]

$$r_{\mathrm{cr}} = \frac{2\alpha}{P' - P} \approx \frac{2\alpha}{n_{\mathrm{a}}(T - T_{\mathrm{melt}})}$$

where α is the surface tension between the liquid and gas, P' is the transient pressure in a gas, and P is the pressure corresponding to the liquid–gas equilibrium: $P = n_{\mathrm{a}} T_{\mathrm{melt}}$. One can consider seed formation as the attachment of a single atom (or monomer) to the seed center, with a characteristic thermal velocity $\nu_{\mathrm{th}} \sim (2T/M)^{1/2}$. This approach allows the estimation of the cluster size under conditions close to those in equilibrium. There have been numerous attempts to improve a classical theory of cluster formation by molecular beams under nonequilibrium conditions without considerable success. Review and discussion of these theories by Milani and Iannotta (1999) is available [3].

The critical radius is directly proportional to the surface tension between the liquid and gas phases and inversely proportional to the pressure of the expanding vapor. Therefore, as one can easily estimate from the above formula, the nanometer size clusters can be formed via the condensation process only in a very dense vapor with a density a few times, 3 to 10 times, lower than the solid density. It means that the formation of nanoclusters by vapor condensation from the supersaturated state might occur in a region very close to the ablated target. Note that the above description of minidroplet formation applies to a monoatomic homogeneous vapor in thermal equilibrium.

Let us now compare the above model to the experimental data of cluster formation. Hot silicon vapor was produced by laser ablation (ArF laser, $\lambda = 193$ nm, 15 ns, 1–3.9 J/cm^2) of a monocrystalline silicon sample [83]. Silicon clusters with a diameter in the range of 1–4 nm were formed through interaction with a flow of pure helium (at pressure from 1 to 4 Torr). The silicon target was used as a substrate where the nanoclusters were collected. Patrone et al. and Marine et al. [83, 84] found that cluster size grows in direct proportion to the increase in laser fluence. However, the increase in laser fluence results in increases in the near surface temperature and pressure of vapors. Thus, in thermodynamics, it should result in a decrease in the critical radius of the nucleus of the new phase and therefore make easier the growth of large droplets. It was also demonstrated that the cluster size

decreases as the distance from the laser focal spot increases. The temperature at the substrate decreases as the distance from the focal spot increases. In accordance with the classical theory of expanding vapor condensation [66, 82], this must result in an increase in the number of critical nuclei for condensation, which is in contradiction to the experimental observation. The vapor condensation model has also been mentioned by Sattler [85], in relation to carbon nanotube formation, as a possibility for forming half-fullerene seeds for the further construction of nanotubes. One can note that condensation leads to the formation of a disordered (liquid) droplet. The complicated internal structure is a result of a microscopic atom-to-atom interaction. One can conclude that the direct application of the thermodynamic approach for explanation of the transient processes of nanocluster formation looks rather questionable. It allows an estimation of an approximate cluster size at best. Let's now consider the kinetic approach to the cluster building process.

Kinetics of Vapor–Gas Interaction The flow of atomic carbon is created by laser ablation near the surface of a target. The carbon cloud expands, carbon atoms collide with the filling gas atoms, exchanging energy, diffusing, and finally forming a mixture of carbon and filling gas with some average transient density and temperature. Thus, the conditions for the formation of clusters—"the primeval soup"—are created. The number density and temperature of this mixture change with the distance from the target. It is quite different near the target and near the substrate, several centimeters away. The processes of collision, diffusion, and atom-to-atom (atom-to-cluster, cluster-to-cluster) attachment can be described qualitatively on the basis of a simple kinetic theory [54, 86–87]. Depending on the masses of ablated atoms and the atoms of the filling gas, the processes of energy exchange will occur at different rates. If the masses of colliding atoms are comparable (for example, carbon flow with an argon fill), the carbon can loose a significant part of its energy even in a single collision. Hence, efficient energy equilibration occurs after several collisions: the carbon vapor is cooling down and the argon gas is heating up. The main processes which contribute to changes in the density of carbon atoms in the cluster formation zone are the following: delivery of atomic carbon by the target ablation, carbon losses due to diffusion out of the formation zone, and carbon consumption in the cluster formation process. On the basis of kinetic considerations, the scenario of carbon cluster formation in a carbon–argon mixture created by the high repetition laser has been suggested. We point out that the main features of this scenario are applicable to the cluster formation of any other element.

Kinetics: Nucleation by Monomer Addition Laser ablation creates an almost continuous inflow of hot carbon atoms and ions, with an average temperature of a few electron volts, into the experimental chamber. The shock wave generated by each pulse rapidly decelerates in the ambient gas atmosphere, and further propagation of hot atoms proceeds by diffusion. Initially the ambient gas is at room temperature (or the temperature of the furnace). The continuous inflow of hot carbons increases the partial density of carbons in the chamber, along with the temperature in the mixture. When the carbon vapor temperature and the number density reach the level where the probability of carbon–carbon attachment becomes significant, the formation of carbonaceous clusters begins. The carbon consumption rate during this formation process significantly exceeds the evaporation rate due to laser ablation. Therefore, the carbon number density rapidly decreases to the value at which the formation process terminates. Thus, the ablation rate, target parameters, pressure, and properties of the ambient gas determine *the formation time* and, accordingly, *the size* of the cluster formed. Continuing laser evaporation leads again to an increase in the carbon number density to a value sufficient to resume the next cycle of cluster formation.

It has been suggested by Gamaly et al. [87] that the cluster formation process is composed of periodic stages of heating and cluster formation, with the time period depending on the initial argon density, the evaporation rate, and on the carbon attachment reaction rate, which in turn is a function of the temperature and density of the atomic carbon. It is clear that the average temperature of the ambient gas depends on the laser repetition rate. For example, in the case of a high repetition rate laser [73], during the short period of cluster formation (in comparison to the heating period), the argon gas does not cool down but maintains a temperature approximately equal to that required for cluster formation. As a result, the average temperature in a carbon–argon mixture appears to be high enough for formation of sp^2 and sp^3 bonds, as experimentally observed.

The maximum number of atoms in a cluster (or the maximum cluster size) from the kinetic viewpoint is directly related to the cluster formation time defined above. The characteristic time for N atomic cluster assembly in an argon–carbon mixture can be estimated under the assumption that the main building process is dominated by single atom attachment to a bigger cluster, taking the attachment cross section to be equal to the geometrical cross section for elastic collisions. This time is directly proportional to the ambient gas density and inversely proportional to the square of the partial density of single carbon atoms. These predictions qualitatively comply with the experimental data that small clusters ($\sim$6 nm, 10^4 atoms) are preferably formed at a low gas pressure of 1 Torr [37, 73], while carbon nanotubes (10^6 atoms) are formed at pressures of 300–500 Torr [27–30, 87]. Therefore, the kinetic approach can qualitatively predict the formation time and the size of the nanocluster. However, kinetic theory, as well as thermodynamics, fails to explain fundamental issues relevant to the internal structure of the nanoclusters and their unusual material properties already observed experimentally, for example, why clusters having a particular number of particles (magic numbers) are more abundant, which features of the formation process are responsible for the intricate structure of a cluster, and what are the reasons for the formation of crystalline or amorphous clusters.

5.1.2. Cluster Annealing and Stabilization

Laser ablation creates a dense homogeneous mixture of hot ablated atoms and atoms of an ambient gas that proved to be the most appropriate medium for the formation of three-dimensional clusters. The homogeneous mixture has high

symmetry—it is almost isotropic. Kinetics and thermodynamics theories can qualitatively explain why the hot atoms can form a cluster of a definite size, like an amorphous snowball, during the many "sticky" collisions. However, it is well known that nanoclusters have different and sometimes very complicated structures. A fundamental question then arises, namely, how does the transition from such a disordered gas phase to the highly symmetric molecule of fullerene [19], or a carbon nanotube [20, 21], or even the more complicated structural unit of carbon nanofoam [37] take place? The most natural answer may be found in the general theory of phase transitions for condensed matter, in the breaking of the initial high symmetry of a homogeneous gas and in the formation of primary seed structures displaying the specific symmetry obtained in the final structure.

Formation of a three-dimensional cluster breaks the initial high symmetry in the cluster formation zone. In the genesis of fullerene, C_{60}, for example, a lower symmetry cluster, pentagon or hexagon, displaying an axis of rotation of a higher order (5 and 6) is created. From the chemical point of view, the substance in the nanocluster formation zone is a random mixture of different low-dimensional carbon clusters (monomers, dimers, trimers, five- and six-membered rings, etc.) formed in the process of stochastic "sticky" collisions of carbon atoms in the carbon-ambient gas mixture. All of these clusters have a lifetime strongly dependent on the temperature. Generally, this lifetime increases with decreasing temperature. Simultaneously, the spatial amplitude of fluctuations decreases when the temperature falls to the critical temperature and below. The critical temperature is the characteristic temperature at which the particular structure begins to stabilize and becomes rigid [88]. The jump in the order parameter (a singularity), which manifests the appearance of a new symmetry, also occurs at this temperature. The free energy of a newly formed cluster at $T < T_c$ must include, in addition to thermal energy, the internal (strain) energy which is absent from a totally disordered state. The configuration entropy of a cluster also characterizes its particular structure. The minimization of the free energy determines the critical temperature for the defect-mediated phase transition from a disordered gas phase [89, 90] or, for the case considered here, the optimum temperature for nanocluster formation. This critical temperature, or rather optimum temperature for nanostructure formation, was found experimentally: for carbon nanotube formation this temperature is approximately 1200–2000 K [29, 40, 43]. For fullerene formation this temperature as seen in molecular dynamics simulations [91] lies in the range of 2000–3000 K.

The formation of three-dimensional carbon clusters can be understood in terms of a phase transition from a disordered to an ordered phase by the introduction of a specific defect assembly or seed with a new structure. In the case of fullerene formation, it is associated with the appearance of pentagon-shaped, or five-wedge, disclinations. This is similar to the conventional scenario seen for disorder–order transition phenomena in the growth of crystalline structures [88, 89]. The appearance of a disclination—a singularity in order parameter—in a hot mixture of unassociated carbon atoms echoes arguments presented by Anderson that "the fluctuations do become of quasi macroscopic size and dimension at the critical point..." [88]. The defect in this context is a means for replacing the symmetry of an initial phase with that of a new one.

We should note that most of the studied three-dimensional carbon clusters display very complicated internal structure. For example, fullerenes and fullerene-like structures are embedded with pentagons, which are units relevant only to surfaces of positive curvature. It is likely that carbon clusters, which include surfaces with negative curvature where heptagons are the responsible topographic elements, have also been identified experimentally [37].

5.2. Electronic and Structural Properties of Individual Nanoclusters

Many intrinsic properties of a particular bulk material, become dependent on the size of a system when the cluster size decreases to several nanometers. The thermodynamic and electrodynamic arguments which apply to the infinite system (when one can ignore boundary effects) become invalid for the description of the properties of very small particles.

The unique features of nanoclusters are directly related to their small size. That size lies between the atomic size and the macroscopic dimensions of many particles in a bulk structure. There are at least two main consequences for the physical and chemical properties resulting from size. First, a nanocluster is a system of many, but a limited, number of particles, say $10–10^4$, and behaves as a large atom with discrete energy levels. Second, accompanying a decrease in cluster size, the ratio of the surface to volume increases. In this case, the surface phenomena dominate the electronic and optical properties of a cluster. For example, in a cluster containing 1000 atoms, about a quarter of the atoms lie close to the surface, suggesting that these atoms may strongly influence the cluster properties.

5.2.1. Critical Cluster Size

It is crucial for understanding the properties of nanoclusters to know at what cluster size (or at what number of atoms in a cluster) the material properties of an atomic system approach those of the bulk structure. Moreover, it is important to understand if all the material properties are changing at the same critical "threshold" size or if different properties change at different cluster sizes.

It was theoretically predicted [92] that small metal clusters ($n < 10$) would have a face-centered cubic (fcc) structure, while in bulk, the structure is body-centered cubic (bcc). It was observed later on [93] that a structural evolution from amorphous to fcc, and subsequently to bcc, structure occurs in tungsten when the number of atoms in a cluster increases, in agreement with Tomanek et al. [92]. The critical cluster size for metallic chromium nanoclusters is determined to be 490 ± 100 atoms, and for Mo nanoclusters it is in the range of 1460–3900 atoms for the cluster–bulk transformation in a structural sense to be completed [94]. The surface-to-volume ratio is huge for small clusters and therefore the surface energy contribution dominates in the total cluster energy. The fcc structure is more compact than the bcc and thus provides a lower surface energy and overall lower total energy

of a cluster. As will be shown in the next section, the critical cluster size appears to be different for different materials and for different properties.

5.2.2. Size Effects

Quantum Size Effects A system of particles in a finite volume has discrete energy levels. In general the distances between energy levels decrease exponentially with increases in the number of particles. In bulk, the interference of electronic wave functions results in a band structure. However, at the opposite limit, when the size of the system decreases to the nanometer scale, the energy bands split again into energy levels. For example, the difference in energy between the electronic levels, $\Delta\varepsilon$, in a small metal cluster increases in inverse proportion to the size of the cluster, R, as $\Delta\varepsilon \sim \nu_F/R$, where ν_F is electron Fermi velocity [95]. If the applied electric field is small in comparison to the interatomic field $eER \ll \Delta\varepsilon$, and the frequency of the external field complies with condition $\omega \ll \nu_F/R$, then the small metallic particle behaves in such a field like an atom with a certain polarizability. A moderately small cluster possesses a larger polarizability than that following from the classical polarization of a metal sphere.

Size-dependent photoluminescence of silicon nanoclusters has been observed experimentally where the cluster size ranged from 1 to 3.5 nm [83, 84]. Analysis of the internal structure of the silicon clusters using high-resolution transmission electron microscopy (HRTEM) revealed that clusters had a crystalline structure and nearly spherical shape [84]. When the cluster sizes were decreased from 3.5 to 1 nm, the peak in the luminescence spectrum shifted from 750 to 300 nm.

Neutral silicon clusters ranging in diameter from 1 to 20 nm were produced by laser ablation of a silicon target in an ambient atmosphere of helium at a pressure of 8 Torr and extracted into a vacuum as a cluster beam. The cluster beam was then scattered by an argon beam flowing in a direction perpendicular to the cluster beam. Clusters of different sizes were deposited onto substrates positioned at different distances from the axis of the cluster beam. On the substrate placed directly on the beam axis, spherical isolated clusters with a mean diameter of 10 nm were deposited. The average size of the clusters decreases as the distance from the axis of the cluster beam increases.

The nanocluster films exhibit a strong red photoluminescence after being exposed to air. The energy of the photoluminescence peak changes between 1.42 eV (845 nm) and 1.72 eV (700 nm) depending on the substrate position: the farther from the cluster beam axis, the higher the energy. Wu et al. [96] relate the observed energy shift to the quantum confinement of carriers in surface-oxidized silicon nanocrystals. A blue shift in the energy of the photoluminescence peak, of up to 2.1 eV, in surface-oxidized porous silicon quantum dots was also observed by another research group when the size of the dot decreased from 3 to 2 nm [97].

These observations qualitatively comply with the effect of increasing the distance between the energy levels while the cluster size decreases. Note that such a luminescence has not been observed in bulk silicon.

Formation Energy It was experimentally determined that the energy for the formation of individual clusters depends on their size and shape. Tin nanoclusters containing from 95 to 975 atoms were formed by laser ablation of tin in a helium atmosphere [98]. For one type of tin clusters it was found that the formation energy was proportional to $(1.64 \pm 0.04 \text{ eV}) \times N^{-1/3}$, indicating compact spherical-like shapes. Another class of clusters had almost constant formation energy of 0.4 ± 0.05 eV. The theory developed for elongated neutral silicon clusters indicates [99, 100] that the surface-to-volume ratio of in these clusters is constant, which indicates that the formation energy was independent of size. It was suggested that the second type of cluster has a quasi-one-dimensional geometry. Thus, in the first type of cluster, the surface energy changes with the change in the number of atoms in the cluster, while in the second cluster type, the contribution of the surface energy to the total cluster energy is the same for clusters of different sizes. The experiments [98] clearly demonstrate the strong influence of surface energy on the formation energy of a cluster: the smaller the cluster is, the easier it is to form.

Melting Temperature: Solid-like to Liquid-like State Transition It has been demonstrated both experimentally and theoretically that the melting temperature of nanoparticles and nanorods is significantly lower than that for the bulk material (see, for example Wang et al., 2002 [101] and references therein). The melting point of platinum nanowire was found to be 400 °C [102] (the bulk value 1772 °C), while the melting temperature of 4.6-nm-thick palladium was 300 °C, which is drastically lower than the bulk value of 1552 °C [103]. Through molecular dynamic simulations [101], the melting temperature of gold helical cylindrical nanowires was found to be ~1100 K, that is, lower than the bulk value of 1357 K, but higher than that of gold nanoclusters.

Usually equilibrium melting starts from the surface and then propagates into the interior. Surface atoms have the fewer nearest neighbors and weaker binding, which may lead to a lower melting temperature at the surface than that for the bulk. Large surface-to-volume ratios and quantum size effects are the two major factors believed to be responsible for this dramatic decrease in the melting temperature of nanoclusters compared to the bulk value.

In contrast, it was found that the interior melting temperature in gold nanorods appeared to be lower than that of the atoms on the surface. Melting starts from the interior atoms, while surface melting occurs at relatively high temperature. This unique thermodynamic behavior, compared to the above description, is closely related to the helical structure of the nanorod. In this case, surface melting is responsible for the overall melting of the structure, and the effect of higher interior melting temperature entirely relates to the complicated internal structure of a rod [101].

The transition from a solid-like to a liquid-like state for finite systems, where the surface and boundary effects are dominant, is not as well defined as that for a bulk solid. The transition may exhibit the intermediate state of the coexistence of different structural phases (isomers) of the same cluster [104]. The solid–liquid phase transition in a system with a finite number of particles is described in terms

of a potential energy surface [105]. Stable clusters correspond to minima on the potential energy surface. If a cluster acquires sufficient energy, the transition from one isomeric form of a cluster to another corresponds to a "jump" over the saddle-like potential barrier separating the minima. The more liquid-like a cluster becomes (the higher the temperature and internal energy), the shorter the time for a transition from one isomeric form to another. One can define the transient state as a state when the lifetimes for the different isomeric forms of a particular cluster become comparable [104].

The lifetime of cesium halide clusters with different spatial arrangements of constituent atoms were observed using temperature-dependent photoelectron spectra and applying the pump-probe technique [104]. The clusters were produced in a laser vaporization source. A plume of laser-ablated cesium halide vapor was swept through a temperature-controlled nozzle by a flow of helium. The clusters were formed through helium–vapor interaction and then mass selected using time-of-flight mass spectrometry. The anions of desired mass were directed into a magnetic bottle, where the electrons were then detached by the action of 1- to 2-ps-long probe laser pulses, and the electron spectra were measured. Three different isomers of cesium halide consisting of four cesium atoms and three iodine atoms were identified: cube, flat ladder, and ring. The transition between these structures proceeds in the following succession: cube $\rightarrow$ ladder $\rightarrow$ ring. At a temperature $\sim$500 K, the lifetime for all three isomers become of the same order of magnitude of several tens of picoseconds. Therefore this temperature can be considered the cluster melting temperature: at this temperature all three isomers are continuously transforming to each other, thus making any particular structure indistinguishable, as in a liquid drop. From the measurements, the time for this phase transition to occur was estimated to be in the range of $\sim$100 ps. One may note that the melting temperature for the bulk cesium halide is 900 K.

Adsorption Energy Two other important microscopic parameters, which depend on the cluster size, are the adsorption energy and the bond distance of single atoms on surfaces. The adsorption energy, or energy of an interaction of an ad-atom and the surface, increases as the number of atoms decreases. Conversely, when the number of atoms in a system increases, this energy converges to the final constant value corresponding to the bulk solid [106]. For example, the adsorption energy for a copper atom on a four-atom copper cluster comprises 3.6 eV, whereas the adsorption energy saturates at 2.7 eV for clusters with a number of atoms larger than 56 [106].

5.2.3. Dielectric Function for a Single Metal Cluster

The qualitative dependence of the dielectric function of a single metal cluster on its size can be understood on the basis of simple arguments of general physics. Let us assume that the dielectric function for the metal cluster in an external electric field with frequency ω has the same form as the function for a conventional metal, for example, having the Drude form [63]:

$$\varepsilon = 1 - \frac{\omega_{\mathrm{pe}}^2}{\omega(\omega + i\nu_{\mathrm{eff}})}$$

here $\omega_{\mathrm{pe}} = (4\pi e^2 n_{\mathrm{e}}/m_{\mathrm{e}})^{1/2}$ is the characteristic electron plasma frequency, n_{e}, m_{e} are the electron number density and electron mass, respectively, and ν_{eff} is the effective frequency of electron collisions with the lattice, which is responsible for the energy dissipation. We assume that the cluster of radius R contains N atoms of a metal with one conductivity electron per atom. One can reasonably suggest that all conductivity electrons fill the whole volume of the cluster homogeneously like a jellium [2, 107, 108]. The electron number density can be expressed as $n_{\mathrm{e}} = Ne/(4/3\pi R^3)$, thus giving the dependence of the plasma frequency on the cluster size.

One can see that the energy dissipation of an external electric field imposed on a small cluster also depends on the size of the cluster. Indeed, in a nanometer size metallic sphere, the mean free path of an electron between collisions significantly exceeds the cluster diameter $l_{\mathrm{mfp}} \gg R$. Therefore, the electron can loose its energy only in collisions with the cluster. The characteristic frequency of this process is proportional to ν_{F}/R. Hence, the effective frequency of energy dissipation is $\nu_{\mathrm{eff}} \sim A\nu_{\mathrm{F}}/R$, where A is a dimensionless proportionality coefficient. It was shown that the classically derived $(1/R)$ law also follows from quantum mechanical calculations (see the thorough discussion of this problem and references by Krebig and Vollmer (1995) [107]). It appears that this law represents a fundamental quantum size effect; it also reflects the surface-to-volume ratio. These simple relations show how the optical properties of metal nanoclusters can be controlled by their size, which suggests many applications. We should note that the dielectric function of a single isolated cluster has not yet been determined experimentally.

5.2.4. Static Polarizability of a Single Metal Cluster

From the viewpoint of classical electrodynamics [63], a metal sphere in an external static electric field E acquires an induced dipole moment $P = R^3E$, which is proportional to the volume of the sphere. However, close consideration of the dipole moment even in frames of a simple jellium model shows that electrons are spilled out of the cluster boundary (defined as $R \sim N^{1/3}$) at a distance δ, depending on the nature of the metal. Usually, this distance is $\delta \sim 1.3$–1.5 Å. Thus, the nanocluster dipole moment is expressed as $P = (R + \delta)^3 E$. These estimations, in the frame of a simple jellium model, are well supported by quantum mechanical calculations [107]. The important conclusion follows that the small metal sphere possesses a polarizability larger than that predicted by classical electrodynamics due to the effective increase in the cluster radius caused by spilling out the electric charge. The change in polarization can be controlled by the size of the particle. This effect was observed experimentally [2] with reasonably good agreement with the theory.

5.2.5. Shell Model for a Simple Metal Cluster

The abundance mass spectra of alkali metal clusters produced and detected by Brack (1993) [108] revealed a striking feature: clusters with a particular mass value, containing a specific number of atoms—"magic numbers"—appeared to be much more abundant in comparison to the average mass of the clusters produced. It is established now that this characteristic abundance pattern holds up even for clusters containing several thousand atoms [2, 108]. It was recognized that the abundance pattern reflects the electronic structure of the cluster and is indicative of the most stable clusters. Clusters of metal atoms are considered a unified edifice having a common electronic structure, with all valence electrons belonging to all constituent atoms and filling corresponding shells that are characterized by specific quantum numbers for such an edifice as a whole. Therefore, maxima in the mass abundance spectra have been identified as clusters with closed electronic shells. Just as for atoms, the electronic system of a cluster with exactly the right number of electrons to complete a shell is very stable. If one more atom is added to a cluster with a closed shell, the valence electrons of this extra atom will occupy the higher energy state, and hence the stability of the new cluster is reduced. The reduced stability is reflected in reduced abundance, which explains the experimentally observed drops in the cluster abundance spectra after each shell-closing number [2]. The theory also predicts that closed-shell clusters have a spherical shape, while open-shell clusters are significantly distorted. Molecular dynamics and quantum chemical calculations later supported the initial simple theories, revealing fundamental physical features of the shell model. The theory of the cluster shell model is now well developed, with many predictions experimentally confirmed (for comprehensive review, see de Heer (1993) and Brack (1993) [2, 108] and references therein). While it is difficult to determine the cluster size and shape experimentally, through measurements of the optical properties of the clusters it is possible to make some conclusions about the influence of shape on the properties of the clusters.

5.2.6. The Affinity between the Internal Structure, Material Properties, and Formation Conditions of Nanoclusters

The discovery and studies of new nanostructured carbon phases, such as fullerenes [19, 51] and nanotubes [20, 21, 56], has opened a new era in materials science. It has been discovered that minute changes in the spatial arrangement of carbon atoms in the space scale of tens of angstroms can profoundly change the electronic properties of these systems from a semiconductor to a metal [110, 111], from a conductor to a superconductor [112, 113], and even from diamagnetic to ferromagnetic [114], with a drastic difference from the bulk properties of graphite, diamond, or amorphous carbon. One could expect that structural rearrangements at the nanometer scale might also change the properties of other materials in ways similar to those with the carbonaceous structures. To date thorough studies have been performed with carbon nanoclusters. Below we concentrate on structural and material properties in connection with the relevant formation conditions of carbonaceous clusters and clustered films.

Laser ablation was used to generate carbonaceous nanoclusters under different experimental conditions: in the ambient gas at different pressures; with additional gas heating; and with different catalysts. A broad variety of different nanostructures have been created which now form the family of "fullerenes." This family includes single-shell fullerenes, C_{60} and C_n [19], multiwalled hollow nanoparticles (fullerene onions), single-wall and multishell carbon nanotubes [20, 21, 29, 56], hypothetical carbonaceous structures with negative curvature [115–118], carbon nanohorns [119], and carbon nanocones [120]. Both the nanocluster formation energy and kinetic factors define the formation conditions and therefore the abundance of the particular nanocluster [118, 121]. The primary product of the interaction between the laser-produced plume and a noble gas is a complicated network of different nanostructures mixed together and interconnected. In order to study the individual nanotubes the meticulous processes of separation and purification of the primary material are used [21]. Another method of cluster selection and separation is to preionize the clusters using a special short-wavelength laser and then to selectively remove clusters with an applied electric field [2, 3, 45, 122].

Cage-like Structures A plume of carbon atoms and ions ejected during the laser ablation of a graphite target was made to interact with a noble gas buffer [45, 122]. Carbon clusters created in such an interaction were then detected by a mass spectrometer. The measured abundance spectra may be regarded as an accumulated "snapshot" of salient features in the nucleation and growth of carbon clusters in a time sequence. A typical mass spectrum of positively charged clusters is bimodal [45, 122]. The peaks in the mass abundance spectra correspond to clusters of enhanced stability. The maximum in the first group of clusters ($n < 30$) corresponds to clusters comprising 11 carbon atoms, while the maximum in the second group ($30 < n < 100$) relates to the famous fullerene, C_{60}, molecule consisting of 60 carbons [19, 122]. Any closed-shell-like exclusively carbon cage is referred to as a fullerene [123]. The fullerenes are viewed as graphitic sheets distorted by inclusions of topological defects, such as five-membered rings, or pentagons. The simplest ball is a pentagonal dodecahedron comprising 20 carbons. The appearance of a pentagon is the signature of a curved surface with positive Gaussian curvature. C_{60}, which is the most stable, and central to the fullerene family, consists of 12 pentagons and 20 hexagons.

The range of clusters is composed of one- and two-dimensional structures (linear and ring clusters, $n = 3$–20) along with three-dimensional clusters, like the closed cage of the fullerene molecule. Multiwalled hollow nanoparticles (nested fullerenes or onions) can be also formed under similar conditions [123]. All these clusters are formed in an almost homogeneous mixture of carbon and noble gas atoms at pressures of several hundred Torr and temperatures of 1000–2000 K. The estimates of the scale of the temperature for fullerene formation may be made from the self-energies of nanoclusters [118, 121]. The additional energy that it is necessary to spend for fullerene formation, in comparison

to the flat graphite sheet, ranges from 0.6 to 0.4 eV/atom [118, 121]. This means that the optimum temperature for the formation of such clusters can be in a range of several thousand degrees Kelvin. This estimate qualitatively complies with the experimental measurements.

The relative population of a particular cluster is governed by cluster growth kinetics. In turn these kinetics depend on the transient partial number density of carbon atoms in the carbon–noble gas mixture and the transient temperature in the cluster formation zone. Therefore, one can conclude that under isotropic conditions, no cluster of preferential symmetry can be formed but only the simple clusters, or clusters with a high symmetry of the order 5 and 6, like fullerenes.

Carbon Nanotubes Another carbon nanocluster, the nanotube, is also produced by laser ablation at high pressure in an almost homogeneous mixture of carbon atoms, noble gas atoms, and a metal catalyst at a temperature of around 1000–2000 K. The required optimal ambient gas pressure was found to be around 500 Torr for the formation of carbon nanotubes, while at 100 Torr the efficiency of carbon nanotube formation was drastically reduced. Carbon nanotube can be imagined as a graphitic sheet rolled up and stitched without seams. Therefore, nanotubes consist entirely of sixfold rings, hexagons. The straight nanotube possesses an axial symmetry, which is apparently absent in the zone of formation. However, it was observed that carbon nanotubes formed in a homogeneous carbon–noble gas mixture were only formed in the presence of metal catalysts at a concentration of ~1% of the total carbon content. These catalysts are usually transitional metals, such as cobalt, iron, or nickel or a mixture of them. Thus, there is strong experimental evidence that a catalyst is responsible for the formation of axially symmetric structures. It may be that catalytic particles are either atoms or metallic clusters containing many atoms; however, there is no thorough theoretical explanation of catalyst action in the formation process.

Nanotubes in general are helical, meaning that the tube is twisted around its axis. In the spirit of defects and energy arguments, one can assume that tubules grow helically with the introduction of a screw dislocation, as metal "whiskers" grow [113]. The theory shows [110, 111] that a chiral tube can be generated in such a way that no distortion of bond angles is introduced. This distortion relates only to the cylindrical curvature of a tube. Helicity is responsible for one of the most striking properties of nanotubes. Depending on chirality one third of the tubules will be metallic and two thirds semiconducting. The existence of metallic and semiconductor nanotubes has been observed experimentally [124]. Raito et al. [111] have suggested that a large fraction of metallic tubules might be formed if the initial seed for the tube cap is centered around a pentagon. However, the proper conditions for the preferential formation of such seeds needs to be formulated. These properties imply many important applications in the future. For example, one can imagine nanometric conducting wires, microscopic metal–semiconductor devices for nanoelectronics or photonics, and many other possibilities. Nanotubes also possess extraordinary strength, with the highest Young modulus of all known materials. This was deduced from the observation of temperature vibrations of freestanding tubes in a transmission electron microscope [125].

It follows from molecular dynamics calculations and a simple elasticity approach [118, 121] that the energy necessary for the formation of tubes consists entirely of the energy required to bend a graphite sheet into a tube, which is ~0.2 eV per carbon atom. Comparison to the fullerene self-energy implies that the tubes are more stable than fullerenes. Another conclusion is that the temperature required for nanotube formation may be somewhat lower than that for the creation of fullerenes and may be around 2000 K, which qualitatively complies with the experimental data presented in Section 2.

Structural Units of Carbon Nanofoam According to theoretical predictions [115–118], hypothetical carbonaceous structures with negative, or hyperbolic, Gaussian curvature, branded "schwarzite," are more complex than fullerenes (elliptic curvature) and buckytubes (parabolic curvature). Sevenfold rings (heptagons) are the topological elements (disclinations) responsible for the generation of schwarzite structures. The exact shape of a single negative disclination is not yet known; however, calculations show that heptagons, in terms of local structure, are energetically less costly (more stable) than pentagons, with the self-energy per carbon in a range of 0.1–0.15 eV/atom [118, 121]. This also means that the temperature conditions for hyperbolic structure formation may be close to those for carbon nanotubes. It has been suggested that the complicated spatial structure of schwarzites might result in unusual electronic properties.

Recently, low-density cluster-assembled carbon nanofoam was produced by laser ablation of glassy carbon in an argon-filled chamber at a pressure range of 0.3–1 Torr [37, 114]. Diffusion-limited aggregation of carbons in the carbon–argon mixture and the subsequent fast quench resulted in a unique, fractal, all-carbon foam deposited on a substrate. Preliminary studies revealed the presence of a hyperbolic schwarzite structure, rather than the expected nanotube-like structures. The foam structure was characterized by scanning electron microscopy, transmission electron microscopy, HRTEM, small-angle X-ray scattering, Raman spectroscopy, EELS, Rutherford backscattering, and surface-area measurements [37, 54, 114]. These studies revealed that the foam is made of a web-like fractal network of randomly interconnected clusters, 6–9 nm in diameter, as can be seen in Figure 8. HRTEM images demonstrated that individual clusters possess an internal periodic structure with a period of ~5.6 Å. The spatial variation of sp^2/sp^3 bonding across a single cluster was mapped using a PEELS with a spatial resolution of 1 nm [37]. These results indicate that sp^2-bonded carbons prevail in the core region of clusters, whereas sp^3-bonded carbons dominate near the cluster boundary. The measured high dc resistivity of the foam suggests that these sp^3-bonded carbon atoms are responsible for sticking the clusters together.

The most salient property of the foam is its unusual magnetic behavior. The freshly produced foam is strongly attracted to a permanent magnet, just like a metallic dust, which demonstrates the existence of an intrinsic positive magnetic moment. Only the structural rearrangement of carbon atoms in the foam can be a reason for the dramatic change in the magnetic properties of this all-carbon system. Electron spin resonance measurements gave the density

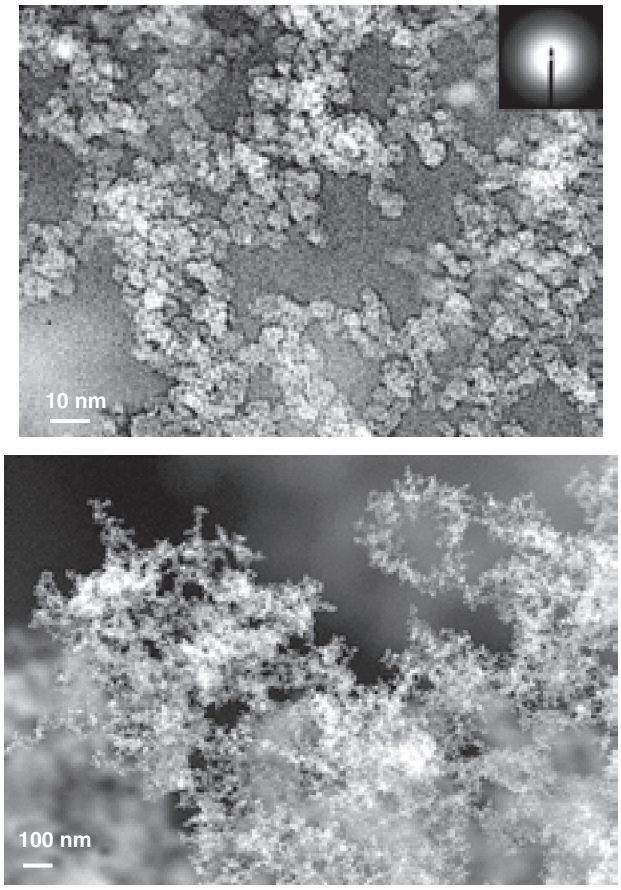


Figure 8. Transmission electron micrograph of carbon nanofoam (left) showing that the foam is assembled from 6-nm clusters. The SEM image (right) shows the web-like structure at lower magnification.

of unpaired spins up to 1.5×10^{20} spins per gram, or one unpaired spin per several hundred carbon atoms.

The ferromagnetic behavior exhibited by the new phase of carbon is extremely unusual in comparison to all known allotropes of carbon. Graphite, diamond, fullerenes, multiwalled carbon nanotubes, and single-walled carbon nanohorns possess a diamagnetic susceptibility in the range of -10^{-5}–10^{-6} emu/g. It seems natural to suggest that the observed ferromagnetic behavior can be traced to the complex microstructure of the foam.

The origin of magnetism in cluster-assembled carbon foam has been theoretically studied using a geometry which contains hyperbolic, negatively curved surfaces [114]. The basic structural unit is a tetrapod that exhibits many of the structural features observed in experiments. The core structure consists of warped sp^2-bonded carbon segments terminated by sp^3 carbon atoms at the hydrogen-passivated edges of the four extremities. In a tetrapod, like in other schwarzite-related structures, tubular segments of 0 Gaussian curvature coexist with convex segments of hyperbolic, saddle-like regions of negative Gaussian curvature. The radius of curvature in this basic unit is close to 6 Å, which is consistent with the superstructure found in the diffraction pattern. The tetrapod-like building blocks can be assembled into a rigid foam structure with a very low density that compares to the measured density of 2–20 $\mu g/cm^3$. The threefold coordinated carbon atoms in the sp^2 regions form a network of hexagons and heptagons only. These trivalent carbons are sterically protected within the system of single and double bonds imposed by the sp^3-terminated tetrapod and occur in groups of three. The number of unpaired spins was found to be robust with respect to size and boundary shape variations within the tetrapod. However, this number depends sensitively on the bonding topology in the regions of the negative Gaussian curvature. Subclusters (tetrapods) containing 264 and 336 carbon atoms have been considered, and the predicted magnetic moment in a C_{264} tetrapod agrees well with the experimental observation.

This is the first time to our knowledge that trivalent carbon radicals embedded into hyperbolic surfaces have been identified in an undoped all-carbon nanocluster prepared under specific conditions by laser ablation. Stabilization of carbon radicals by steric protection has been known since the synthesis of triphenylmethyl by Gomberg in 1900 [126].

The ferromagnetism found in these complex nanoclusters occurs as a consequence of nanometer-sized conducting segments containing atoms with different electronic configurations. It appears that a careful preparation of conditions for nanocluster formation by laser ablation may lead to similar changes in the structural and electronic properties of different materials.

Fractal Nanoclusters Nucleation of a new phase is one of the processes where the fractal nature of the object often manifests itself. The generally accepted models for cluster growth include diffusion-limited aggregation, cluster–cluster assembling, homogeneous nucleation by monomer addition, and others. All the models assume that the nucleus, the seed of the new phase, has a spherical shape on which the further assembling of the structure occurs. Experimental evidence has been presented by Zenkevich et al. [127] which shows that gold clusters of ~5 nm in diameter have a fractal structure with fractal dimension of $D = 1.33 \pm 0.08$. These clusters were deposited on the surface of highly oriented pyrolitic graphite by laser ablation using an Nd:YAG laser with the following parameters of ablation: $\lambda = 1.06$ μm; 15-ns pulse duration; 25-Hz repetition rate; average intensity $\sim 10^9$ W/cm. The authors [127] suggested that the shape and dimension of the clusters are determined by the rate at which atoms arrive at the surface where the formation occurs and by the symmetry of this surface. It was proposed that the formation of clusters proceeds under highly nonequilibrium conditions by the attachment of an ad-atom to the fractal nucleus.

6. CLUSTER-ASSEMBLED MATERIALS

With the laser technology that is currently available, it is possible to ablate any existing material—refractory metals, complex composites, and dielectrics—if the laser parameters are chosen according to the properties of the material. The proper choice of laser parameters and conditions in the experimental chamber (pressure of the ambient gas or vacuum) defines the mass distribution in a cluster flow. The size of the clusters produced ranges from several atoms to tens of thousands of atoms per cluster. Clusters then can be deposited on a substrate to form a cluster-assembled film [3].

Films of clusters embedded in various codeposited matrices have been produced [128]. Very often such a film possesses material properties drastically different from those in the bulk of the initial material.

Cluster networks can be classified into two categories [5]. The first one comprises clusters embedded in a host matrix and well separated from each other. The properties of such structures depend on the properties of the individual clusters and that of the host matrix. The dielectric properties of the cluster-embedded composites depend on low and easy-to-change filling factors (~0.01 vol fraction of clusters) and on cluster-to-cluster contacts and connections. The second category includes cluster networks—sponge, froth, or foam-like structures—with properties which are drastically different from those of isolated clusters and which strongly depend on intercluster connections and interactions.

6.1. Cluster-Assembled Films

The flow of nanoclusters hits a substrate and starts building a deposit. First of all we shall distinguish between cases when the cluster-assembled film grows on a substrate in a vacuum [3, 32, 33] and in a noble gas-filled chamber [37, 83, 84, 114].

Molecular dynamics simulations of film growth by energetic cluster impact in a vacuum have shown the influence of cluster energy on the morphology of the film [129]. It is now widely accepted that the energy of a cluster must be sufficiently low in order to prevent the splitting of a cluster during cluster–substrate impact. The straightforward idea of "random paving" on a substrate seems to be in contradiction to experiments: it appears that film morphology depends on the cluster size [128]. Indeed, gold, antimony, and silver clusters, comprising 250–300 atoms, 2 nm in diameter, form large branched and well-separated islands with definite long-range order when deposited on a graphite surface.

The presence of an evaporation cell in the deposition chamber allows the production of films with clusters embedded into different matrices. Harbich et al. [130, 131] produced films of mass-selected metal clusters (Ag_n and Au_n) embedded in various solid matrices of Ar and Kr using this technique.

Gold clusters, comprising 250–300 atoms, embedded in various matrices (SiO_x, LiF, and MgF_2) are randomly distributed over a film surface. It was demonstrated that the cluster volume fraction of Au clusters embedded into a SiO_x matrix can be easily changed from 2 to 11% [128]. Therefore, films with variable dielectric function can be produced for a range of optical applications.

Another interesting application relates to the deposition of transitional metal clusters onto different substrates to produce magnetic films for high-density memory devices and spin electronics. Films with the size-controlled clusters of iron, cobalt, and nickel have recently been produced [128]. The incident free cluster size distributions were centered on Fe_{150}, Co_{300}, and Ni_{300}. Cluster-assembled films with thicknesses up to 100 nm were produced. These films have a granular structure with a grain size of 3–5 nm, slightly larger than the size of the incident cluster size resulting from the diffusion and coalescence of clusters.

Cluster-assembled films about 100 nm thick were produced with carbon cluster beams. The mass distribution in the beams was centered on C_{20}, C_{60}, and C_{900}. These films have a highly porous nanostructured morphology resulting from a random stacking of incident clusters. The density of the films is around 1 g/cm^3, which is much lower than that of conventional carbonaceous structures. The most interesting structure appears to be C_{20} "amorphous" film with a short-range order of 15 nm, with pure sp^3 hybridization and the absence of any long-range order.

Similar films were produced with silicon cluster beams with the size distribution centered on Si_{50}. It is suggested that the presence of a large number of pentagons may significantly change the electronic structure of the clusters. These silicon films also have a granular structure and they are highly porous. The most unusual feature of these films is a strong visible luminescence at ~750 nm comparable to that seen in porous silicon and recently observed in 3.5-nm silicon clusters [83, 84].

6.2. Magnetic Properties of Cluster-Assembled Films

Monodispersed cobalt nanoclusters, with cluster sizes varying from 300 to 9000 atoms, were produced by the cluster-beam technique [132]. These clusters were embedded in Cu and SiO_2 substrates, where the cobalt volume concentration varied from 10 to 50%. The magnetization of cobalt in such structures is always lower than the bulk value. This magnetization increases with increasing cluster size and decreases with increasing Co concentration for a given cluster size. Calculations of intercluster exchange interactions are used to qualitatively explain the magnetization data as a function of Co concentration in good agreement with the experimental results [132].

6.3. Nonlinear Optical Properties of Cluster-Assembled Films

It was found [133] that Si films deposited by laser ablation exhibit a nonlinear refractive index change, as high as $\Delta n = -0.5$ at a wavelength of $\lambda = 532$ nm for films with an average thickness of 200 nm. These films consist of large droplets composed of crystallites with hexagonal wurtzite symmetry and with nanoclusters interspersed between them. The crystallographic symmetry of these droplets was observed with Raman spectroscopy, as well as linear and nonlinear optical measurements, when the films were annealed under various conditions. The authors attribute the large nonlinear refraction coefficient to the hexagonal wurtzite symmetry of the crystallites, which raises the possibility of developing very efficient nonlinear optical devices [133]. However, the origin of this nonlinearity is not fully understood.

6.4. Electrical Conductivity of Nanoclusters and Cluster-Assembled Materials: Quantum Charge Transport and Localization

Another interesting property of the nanocluster-assembled system relates to electrical conductivity. It appears that the "conductivity quantum," or conductivity scale of $2e^2/h$, plays

an important role in the electric conductance through the "bottle necks" connecting the nanoclusters in the cluster-assembled systems. Simple scaling illustrates the appearance of the "conductivity quantum" or "resistance quantum," h/e^2, when one considers the charge transport at the atomic space and time scales. The motion of electrons with charge e, mass m, and number density n in an applied electric field E of high-frequency ω generates the current density

$$j = en\nu \approx \frac{e^2 n}{m(v + i\omega)} E \equiv \sigma E$$

This is Ohm's law, where v stands for the electron collision frequency responsible for the energy dissipation (or resistance $1/\sigma$). The full current, $I = jS$, flowing through the sample with a cross section S, and length L, and containing a total number of electrons $N = nSL$, can be presented in the form

$$I \approx \frac{e^2 N}{mvL} E \approx \frac{e^2 N}{mvL^2} V$$

We assume $\omega \ll v$, which always holds for metals, and $E \propto V/L$ (where V is a potential). The resistance R, can be expressed as

$$R \approx \frac{NV}{I} = \frac{mvL^2}{e^2}$$

Considering all parameters at atomic scales, that is, taking the collision frequency as the same order of magnitude as the atomic frequency, $\nu \cong e^2/a_B$, and $L \cong a_B$, where $a_B = \hbar^2/me^2$ is the Bohr radius, one obtains the quantum unit of resistance as

$$R_0 \approx \frac{\hbar}{e^2}$$

This unit comprises 2.5812×10^4 ohm. Most probably this value of conductance (resistance) separates two different regimes of high and poor conductivity, giving the percolation and localization regimes [134].

The conductivity in nanowires of atomic dimension behaves in the jump-like manner of quantum charge transport. The electrical conductivity of small silicon clusters ($n =$ 1–10, 13, 20) placed between atomistic aluminium and gold leads has been investigated using the *ab initio* nonequilibrium Green's function formalism [135]. All clusters display metallic conductance ranging between one and two quantum units, $2e^2/h$. The transport properties of these cluster junctions may be understood in terms of both the band structure of electrodes and the electronic states of the cluster, modified by the lead environment and size effects. The resistance in this system is $R \sim R_0$.

Carbon nanofoam composed from well-defined carbon clusters of 6–9 nm in diameter, randomly interconnected in a web-like fractal structure, demonstrates a different limit case for conductivity [49]. The resistivity of the as-deposited foam, measured in the voltage range ± 100 V, demonstrates nonlinear current–voltage characteristics with strong hysteresis. The resistivity of the foam after annealing is equal to $1–3 \times 10^9$ Ω/cm at room temperature and $1–10 \times 10^{13}$ Ω/cm at 80 K, which is similar to that of amorphous diamond-like films. Thus, resistance of the foam exceeds R_0 by many orders of magnitude. One can assume that the electrons in a foam are strongly localized in nanoclusters, and the electron conductivity has a tunneling character.

6.5. Photoluminescence in GaAs Nanoaggregates

Laser ablation of a single crystal GaAs target in a vacuum or Ar gas has produced nanoclusters of GaAs [136]. Atomic force and transmission electron microscopy have shown that most of the clusters were spherical, with diameters in the range of 1 to 50 nm, with a peak size distribution between 5 and 9 nm, depending on the Ar gas pressure or laser fluence. X-ray diffraction, solid state nuclear magnetic resonance, Auger electron spectroscopy, electron energy-loss spectroscopy, and high-resolution transmission electron microscopy revealed that these nanoclusters were randomly oriented GaAs crystallites. An oxide outer shell of $\sim$2 nm subsequently developed on the surfaces of the nanocrystals as a result of transportation of the cluster in air. Unpassivated GaAs nanoclusters exhibited no detectable photoluminescence. After surface passivation, these nanoclusters displayed photoluminescence energies less than that of the bulk GaAs from which they were made. These experiments suggest an abundance of sub-bandgap surface states in GaAs nanocrystals. Thus, this kind of nanocluster with a surface layer demonstrates that a decrease as well as an increase can be achieved with transition from the bulk to the nanoscale level. Such changes in the photoluminescence energy were similar to that observed with silicon clusters [83, 84].

7. PERSPECTIVES ON LASER ABLATION FOR CONTROLLABLE PRODUCTION OF NANOCLUSTERS

Laser ablation has proven to be an efficient and flexible tool for the production of a large variety of novel nanoclusters with remarkable properties. These novel nanostructures include fullerenes, carbon and boron nitride nanotubes, magnetic nanofoam, metal clusters, silicon clusters, and a variety of cluster-assembled films and nanofilms. It is clear that this is only the beginning; many different and complex clusters can be created and tested for future applications. Generation of nanoclusters by laser ablation is already a well-controlled process where the ablation rate, temperature and ionization states of the ablated atomic flux, and the conditions of the laser plume and ambient gas interaction can be controlled with high precision. Moreover, it is now clear that control over these processes in time and space can be significantly improved by the use of short laser pulses, as short as a few femtoseconds, and high repetition rate (up to hundreds of megaHertz) lasers. There are also some obvious extensions for the use of laser ablation for simultaneous coevaporation of several different targets for deposition of complex films or preparation of unusual composites and alloys. The same technique can be used for the preparation of multilayered systems consisting of nanometer-thick layers of different materials, which may be regarded as a one-dimensional analog of cluster-assembled films.

Nanoscience and nanotechnology are still in their infancy. However, one can easily foresee many short-term and long-term applications for nanoclusters and cluster-assembled structures.

The obvious short-term applications for cluster networks are in catalytic devices, systems with unusual and variable dielectric properties, and layers with controllable heat conduction. However, there will be also many long-term applications for nanoclusters, for example, as part of future nanodevices for nanoelectronics, nanophotonics, and spintronics. We may also mention nanowires of carbon nanotubes, nanometric metal–dielectric junctions, and ultra-low-capacitance devices comprising one or several nanoclusters, Coulomb-blockade structures where one-by-one passage of a single electric charge between two neighboring particles can be regulated.

The pulse duration, the intensity on the target surface, and the repetition rate can be precisely controlled in laser ablation, so precise and improved control over the ablated vapor can be achieved. Using lasers one can control heating and annealing rates on a space scale of micrometers and a time scale of femtoseconds. This improvement in the control of cluster formation conditions leads to realization of the atom-to-atom attachment mode of nanocluster building, increasing the possibilities for control over the cluster size and the internal structure of a cluster. Laser ablation is also a route for creating metastable allotropes of known materials where the unusual internal structure determines new material properties such as were seen in conducting carbon nanotubes or paramagnetic all-carbon nanofoam.

Laser ablation has demonstrated the highest deposition rate in the production of thin films, of $\sim$10 μm/h [73, 75], and the highest rate of production of a nanofoam of 1 cm^3/min. If lasers can be made cost effective and manageable under industrial conditions, then the industrial production of nanostructures with the use of lasers is a distinct possibility in the near future.

Indeed, there is still a "plenty of room at the bottom" and lasers will help to fill it with new and amazing nanostructures.

GLOSSARY

Ad-atom An additional atom attached to a nanocluster.

Adsorption Process of attachment of ad-atom to a structure.

Atomized flow Plume or cloud of atoms created by laser ablation.

Binding energy The energy necessary to separate atoms from a molecule. It equals to the energy of vaporisation per atom.

Critical cluster size The cluster size (or critical number of atoms in a cluster) at which the material properties of a nanocluster approach to those for the bulk structure.

Femtosecond, picosecond, nanosecond Time units, respectively, 10^{-15} s, 10^{-12} s, 10^{-9} s, used to characterise fast laser-matter interaction processes such as laser ablation and nanocluster formation.

Laser Abbreviation of the "light amplification by stimulated emission of radiation." Modern lasers are capable of transformation of any of existing materials to atomic state.

Laser ablation Removal of material from a surface by means of laser irradiation. The term 'laser ablation' is used to emphasize the non-equilibrium vapour/plasma conditions created at the surface by intense laser pulse, to distinguish from 'laser evaporation,' which is heating and evaporation of material in conditions of thermodynamic equilibrium.

Nanoclusters, or Nanoparticles Aggregates of atoms or molecules containing between a few and a few thousand atoms that have properties drastically different from those in bulk. The size of nanoclusters is in the range of 1–100 nm (1 nanometer = 10^{-9} m). Nanoparticle behaves as a large atom with discrete energy levels. With a decrease in cluster size the ratio of the surface to volume increases, and the surface phenomena dominate the electronic and optical properties of a cluster.

Nanocluster-assembled film Films formed by clusters of desirable size and mass deposited on a substrate, or film of clusters embedded in various co-deposited matrices.

Nucleation An initial process of cluster growth (formation of nucleus or seed) inside an atomic mixture by aggregation of atoms and clusters.

Quantum charge transport The jump-like behaviour of charge transport between the nanoclusters connected by a "bottle neck" that contains a few atoms. It can be associated with a motion of a few single charges.

Quantum-size effects, quantum confinement The changes in electronic and structural properties as the size of the system decreases to the nanometer scale where quantum effects become dominant—the energy bands split into energy levels.

Self-assembly Formation of nanostructures from an atomic mixture where atoms bond due to inherent tendency of molecules and molecular clusters to interact and organize themselves into structures.

Shell model for cluster electronic structure The cluster is a unified edifice (somewhat similar to a big "atom") having electronic structure with all valence electrons filling the common "shells." Filled shells correspond to most stable clusters. Maximums in the mass abundance spectra correspond to the clusters having the closed electronic shells.

"Sticky" collision Inelastic atomic collision resulting in bonding of atoms.

ACKNOWLEDGMENT

The authors acknowledge Dr. D. Golberg for providing HRTEM photographs of single wall carbon nanotubes for the present review. The authors are grateful to S. T. Hyde, A. G. Christy, and D. Tomanek for many useful discussions, and to K. Gamaly for careful proofreading of the manuscript. E. G. G. acknowledges the assistance of the Australian Research Council under the ARC Centers of Excellence program through the Centre for Ultrahigh bandwidth Devices for Optical Systems (CUDOS).

REFERENCES

1. Klaus Sattler, Ed., "Cluster Assembled Materials," Material Science Forum, Vol. 232. Trans Tech Publications Ltd., Switzerland 1996.
2. W. A. de Heer, The physics of simple metal clusters: Experimental aspects and simple models, *Rev. Mod. Phys.* 85, 611 (1993).
3. P. Milani and S. Iannotta, "Cluster Beam Synthesis of Nanostructured Materials." Springer-Verlag, Berlin, 1999.
4. A. H. Pfund, *Rev. Sci. Instrum*. 1, 397 (1930).
5. P. Marquardt, *Appl. Phys. A* 68, 211 (1999).
6. D. B. Chrisey and G. K. Hubler, Eds., "Pulsed Laser Deposition of Thin Films." Wiley, New York, 1994.
7. P. Caricato, G. Leggieri, A. Luches, A. Perrone, E. Gyorgy, I. N. Mihailescu, M. Popescu, G. Barucca, P. Mengucci, J. Zemek, and M. Trchova, *Thin Solid Films* 307, 54 (1997).
8. Rossi, B. Andre, A. van Ven, P. E. Mijnarends, H. Schut, F. Labohm, M. P. Delplancke, H. Dunlop, and E. Anger, *Thin Solid Films* 253, 85 (1994).
9. C. Vivien, J. Hermann, A. Perrone, C. Boulmer-Leborgne, and A. Luches, *J. Phys. D: Appl. Phys*. 31, 1263 (1998).
10. Y. Nakata, H. Kaibara, T. Okada, and M. Maeda, *J. Appl. Phys.* 80, 2458 (1996).
11. H. F. Sakeek, T. Morrow, W. G. Graham, and D. G. Walmsley, *Appl. Phys. Lett.* 59, 3631 (1991).
12. T. J. Goodwin, V. J. Leppert, S. H. Risbud, I. M. Kennedy, and H. W. H. Lee, *Appl. Phys. Lett.* 70, 3122 (1997).
13. P. Verardi, M. Dinescu, F. Graciun, and A. Perrone, *Appl. Surf. Sci.* 127–129, 457 (1998).
14. S. Gill, A. A. Anderson, R. W. Eason, T. J. Warburton, and D. P. Shepherd, *Appl. Phys. Lett.* 69, 10 (1996).
15. S. Fukaya, T. Hasegawa, Y. Ishida, and M. Obara, High-power laser ablation, in "SPIE Proceedings, Advanced High-Power Lasers and Applications AHPLA'99," 3885-42, Osaka, Japan, 1–5 Nov., 1999.
16. V. E. Bondybey and J. H. English, *J. Chem. Phys.* 76, 2165 (1982).
17. S. J. Riley, E. K. Parks, C. R. Mac, L. G. Pobo, and S. Wexler, *J. Phys. Chem*. 86, 391 (1982).
18. Smalley, R. E., *Laser Chem.* 2, 167 (1983).
19. H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl, and R. E. Smalley, *Nature* 318, 162 (1985).
20. S. Iijima, *Nature* 354, 56 (1991).
21. T. W. Ebbesen, Ed., "Carbon Nanotubes: Preparation and Properties." CRC Press, Boca Raton, 1996.
22. D. Golberg, Y. Bando, M. Eremets, K. Takemura, K. Kurashima, and H. Yusa, *Appl. Phys. Lett.* 69, 2045 (1996).
23. P. Calderola and H. Knoepfel, Eds., "Physics of High Energy Density," Proceedings of the International School of Physics "Enrico Fermi" Course XLVIII. Academic Press, New York, 1971.
24. S. I. Anisimov, Y. A. Imas, G. S. Romanov, and Y. V. Khodyko, "Action of High-Power Radiation on Metals." Consult. Bureau, Springfield, 1971.
25. J. F. Ready, "Effects of High Power Laser Radiation." Academic Press, New York, 1971.
26. R. L. McCrory, Laser-driven ICF experiments, Chap. 22 in "Nuclear Fusion by Inertial Confinement" (G. Velarde, Y. Ronen, and J. M. Martinez-Val, Eds.). CRC Press, Boca Raton, 1993.
27. T. Guo, P. Nikolaev, A. Thess, D. T. Colbert, and R. E. Smalley, *Chem. Phys. Lett.* 243, 49 (1995).
28. W. K. Maser, E. Muñoz, A. M. Benito, M. T. Martínez, G. F. de la Fuente, Y. Maniette, E. Anglaret, and J.-L. Sauvajol, *Chem. Phys. Lett.* 292, 587 (1998).
29. F. Kokai, K. Takahashi, M. Yudasaka, R. Yamada, T. Ichibashi, and S. Iijima, *J. Phys. Chem. B* 103, 4346 (1999).
30. M. Yudasaka, F. Kokai, K. Takahashi, R. Yamada, N. Sensui, T. Ichibashi, and S. Iijima, *J. Phys. Chem. B* 103, 3576 (1999).
31. E. G. Gamaly, A. V. Rode, and B. Luther-Davies, *J. Appl. Phys.* 85, 4213 (1999).
32. A. V. Rode, B. Luther-Davies, and E. G. Gamaly, *J. Appl. Phys.* 85, 4222 (1999).
33. M. D. Perry, B. C. Stuart, P. S. Banks, M. D. Feit, V. Yanovsky, and A. M. Rubenchik, *J. Appl. Phys.* 85, 6803 (1999).
34. A. V. Rode, A. Zakery, M. Samoc, E. G. Gamaly, and B. Luther-Davies, *Appl. Surf. Sci.* 197–198, 481 (2002).
35. A. V. Rode, M. Samoc, B. Luther-Davies, E. G. Gamaly, K. F. MacDonald, and N. I. Zheludev, *Optics Lett.* 26, 441 (2001).
36. E. G. Gamaly, A. V. Rode, O. Uteza, M. Samoc, and B. Luther-Davies, *Appl. Surf. Sci.* 197–198, 730 (2002).
37. A. V. Rode, S. T. Hyde, E. G. Gamaly, R. G. Elliman, D. R. McKenzie, and S. Bulcock, *Appl. Phys. A* 69 (Suppl.), S755 (1999).
38. V. Z. Kolev, M. J. Lederer, B. Luther-Davies, and A. Rode, *Optics Letters* 28, 1275 (2003).
39. A. V. Rode, E. G. Gamaly, and B. Luther-Davies, *Appl. Phys. A* 70, 135 (2000).
40. S. Arepalli, P. Nikolaev, W. Holmes, and C. D. Scott, *Appl. Phys. A* 70, 125 (2000).
41. W. Kautek, S. Pentzien, A. Conradi, J. Krüger, and K.-W. Brzezinka, *Appl. Surf. Sci.* 106, 158 (1996).
42. E. G. Gamaly, A. V. Rode, A. Perrone, and A. Zocco, Mechanisms of ablation rate decrease in multiple pulse laser ablation, *Appl. Phys. A* 73, 143 (2001).
43. A. A. Puretzky, H. Schittenhelm, Xudong Fan, M. J. Lance, L. R. Allard, Jr., and D. Geohegan, *Phys. Rev. B* 65, 245, 425 (2002).
44. E. G. Gamaly, A. V. Rode, and B. Luther-Davies, *Appl. Phys. A* 69(Suppl.), S121 (1999).
45. E. A. Rohlfing, D. M. Cox, and A. Kaldor, *J. Chem. Phys.* 81, 3322 (1984).
46. A. A. Puretzky, D. B. Geohegan, H. Schittenhelm, Xudong Fan, and M. A. Guillorn, *Appl. Surf. Sci.* 197–198, 552 (2002).
47. M. Born and E. Wolf, "Principles of Optics." Pergamon Press, Oxford, 1964.
48. E. G. Gamaly, A. V. Rode, W. K. Maser, E. Munoz, A. M. Benito, M. T. Martinez, and G. F. de la Fuente, *Appl. Phys. A* 70, 161 (2000).
49. M. Ge and K. Sattler, *Appl. Phys. Lett.* 65, 2284 (1994).
50. J. W. G. Wildoer, L. C. Venema, A. G. Rinzler, and R. E. Smalley, *Nature* 391, 59 (1998).
51. D. Golberg, Y. Bando, L. Bourgeois, and K. Kurashima, *Carbon* 37, 1858 (1999).
52. E. G. Gerstner, P. B. Lukins, D. R. McKenzie, and D. G. McCulloch, *Phys. Rev. B* 54, 14504 (1996).
53. J. Bulir, M. Jelinek, V. Vorlicek, D. Chvostova, and L. Soukup, *J. Non-Cryst. Solids* 188, 118 (1995).
54. A. V. Rode, E. G. Gamaly, and B. Luther-Davies, *Appl. Phys. A* 70, 135 (2000).
55. A. M. Rao, E. Richter, Shunji Bandow, Bruce Chase, P. C. Eklund, K. A. Williams, S. Fang, K. R. Subbaswamy, M. Menon, A. Thess, R. E. Smalley, G. Dresselhaus, and M. S. Dresselhaus, *Science* 275, 187 (1997).
56. M. S. Dresselhaus, G. Dresselhaus, and P. C. Eklund, "Science of Fullerenes and Carbon Nanotubes." Academic Press, New York, 1996.
57. R. F. Egerton, "Electron Energy-Loss Spectroscopy in the Electron Microscope." Plenum, New York, 1986.
58. S. D. Berger, D. R. McKenzie, and P. S. Martin, *Philos. Mag. Lett*. 57, 285 (1988).
59. S. Brunauer, P. H. Emmett, and E. Teller, *J. Phys. Chem.* 60, 309 (1938).
60. W. D. Harkins and G. Jura, *J. Chem. Phys.* 11, 431 (1942).
61. P. W. Schmidt, in "Modern Aspects of Small-Angle Scattering" (H. Brumberger, Ed.). Kluwer Academic, Dordrecht, 1995.
62. J. A. Leavitt, L. C. McIntyre, Jr., and M. R. Weller, Backscattering spectrometry, in "Handbook of Modern Ion Beam Material

Analysis" (J. R. Tesmer and M. Nastasi, Eds.), pp. 37–81. MRS, Pittsburgh, 1995.
63. L. D. Landau and E. M. Lifshitz, "Electrodynamics of Continuous Media." Pergamon Press, Oxford, 1960.
64. W. L. Kruer, "The Physics of Laser Plasma Interaction." Addison–Wesley, New York, 1987.
65. E. G. Gamaly, A. V. Rode', and V. T. Tikhonchuk, *Phys. Plasmas* 9, 949 (2002).
66. Y. B. Zel'dovich and Y. P. Raizer, "Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena." Academic Press, New York, 1967.
67. B. C. Stuart, M. D. Feit, S. Herman, A. M. Rubenchik, B. W. Shore, and M. D. Perry, *JOSA B* 13, 459 (1996).
68. P. P. Pronko, P. A. Van Rompay, C. Horvath, F. Loesel, T. Juhasz, X. Liu, and G. Mourou, *Phys. Rev. B* 58, 2387 (1998-I).
69. P. A. Van Rompay, M. Nantel, and P. P. Pronko, *Surf. Coatings Technol.* 100–101, 496 (1998).
70. D. Bauerle, "Laser Processing and Chemistry." Springer-Verlag, Berlin, 2000.
71. Experimental Methods in the Physical Sciences, Vol. 30, in "Laser Ablation and Desorption" (J. C. Miller and R. F. Haglund, Jr., Eds.). Academic Press, New York, 1998.
72. A. A. Voevodin and M. S. Donley, *Surf. Coatings Technol.* 82, 199 (1996).
73. E. G. Gamaly, A. V. Rode, and B. Luther-Davies, *Laser Particles Beams* 18, 245 (2000).
74. P. Pronko, P. A. Van Rompay, and S. Sato, "Proceedings of SPIE," Vol. 3269, 1998.
75. P. S. Banks, L. Donh, B. C. Stuart, M. D. Feit, A. M. Rubenchik, A. M. Komashko, M. D. Perry, and W. McLean, *Appl. Phys. A* 69, 347 (1999).
76. J. N. Leboeuf, K. R. Chen, J. M. Donato, D. B. Geohegan, C. L. Liu, A. A. Puretzky, and R. F. Wood, *Phys. Plasmas* 3, 2203 (1996).
77. M. P. Siegal, L. J. Martinez-Miranda, N. J. DiNardo, D. R. Tallant, J. C. Barbour, and P. M. Provencio, High Power Laser Ablation, "Proceedings of SPIE," Santa Fe, NM, 1998.
78. O. Eibl and B. Roas, *J. Mater. Res.* 5, 2620 (1990).
79. P. Schwab and D. Bauerle, *Physica C* 182, 103 (1991).
80. D. P. Norton and D. H. Lowndes, *Appl. Phys. Lett.* 63, 1432 (1993).
81. P. Gonzalez, R. Soto, E. G. Parada, X. Redondas, S. Chiussi, J. Serra, J. Pou, B. Leon, and M. Perez-Amor, *Appl. Surf. Sci.* 109–110, 380 (1997).
82. L. D. Landau and E. M. Lifshitz, "Statistical Physics." Pergamon Press, Oxford, 1980.
83. L. Patrone, D. Nelson, V. Safarov, M. Sentis, and W. Marine, *J. Luminescence* 80, 217 (1999).
84. W. Marine, L. Patrone, B. Luk'yanchuk, and M. Sentis, *Appl. Surf. Sci.* 154–155, 345 (2000).
85. M. Ge and K. Sattler, *Science* 260, 515 (1993).
86. E. G. Gamaly and L. T. Chadderton, *Proc. R. Soc. London Ser. A* 449, 381 (1995).
87. E. G. Gamaly, A. V. Rode, W. K. Maser, E. Munoz, A. M. Benito, M. T. Martinez, and G. F. de la Fuente, *Appl. Phys. A* 70, 161 (2000).
88. P. W. Anderson, "Basic Notions of Condensed Matter Physics." Benjamin/Cummings, Menlo Park, CA, 1984.
89. J. M. Kosterlitz and D. J. Thouless, *J. Phys. C: Solid State Phys.* 5, L124 (1972).
90. H. S. Seung and D. R. Nelson, *Phys. Rev. A* 38, 1005 (1988).
91. S. G. Kim and D. Tomanek, *Phys. Rev. Lett.* 72, 2418 (1994).
92. D. Tomanek, S. Mukherjee, and K. H. Benneman, *Phys. Rev. B* 28, 665 (1983).
93. S. J. Oh, S. H. Huh, H. K. Kim, J. W. Park, and G. H. Lee, *J. Chem. Phys.* 111, 7402 (1999).
94. S. H. Huh, H. K. Kim, J. W. Park, and G. H. Lee, *Phys. Rev. B* 62, 2937 (2000-II).
95. L. P. Gor'kov and G. M. Eliashberg, *Sov. Phys. JETP* 21, 940 (1965).
96. H.-P. Wu, A. Okano, and K. Takayanagi, *Appl. Phys. A* 71, 643 (2000).
97. M. V. Wolkin, J. Jorne, P. M. Fauchet, G. Allan, and C. Delerue, *Phys. Rev. Lett.* 82, 197 (1999).
98. T. Bachels, R. Schafer, and H.-J. Guntherodt, *Phys. Rev. Lett.* 84, 4890 (2000).
99. E. Kaxiras and K. Jakson, *Phys. Rev. Lett.* 71, 727 (1993).
100. J. C. Grossman and L. Mitas, *Phys. Rev. B* 52, 16735 (1995).
101. J. Wang, X. Chen, G. Wang, B. Wang, Wei Lu, and J. Zhao, *Phys. Rev. B* 66, 085408 (2002).
102. Z. Liu, Y. Sakamoto, T. Ohsuma, K. Higara, O. Terasaki, C. H. Ko, H. J. Shin, and R. Ryoo, *Angew. Chem. Int. Ed. Engl.* 39, 3107 (2000).
103. K. B. Lee, S. M. Lee, and J. Cheon, *Adv. Mater.* 13, 517 (2001).
104. A. J. Dally and L. A. Bloomfield, *Phys. Rev. Lett.* 90, 063401 (2003).
105. J. P. Rose and R. S. Berry, *J. Chem. Phys.* 96, 517 (1992).
106. T. Jacob, S. Fritzsche, W.-D. Sepp, B. Fricke, and J. Anton, *Phys. Lett. A* 300, 71 (2002).
107. U. Krebig and M. Vollmer, "Optical Properties of Metal Clusters." Springer-Verlag, Berlin, 1995.
108. M. Brack, *Rev. Mod. Phys.* 65, 677 (1993).
109. R. Gao, Z. L. Wang, Z. Bai, W. A. de Heer, L. Dai, and M. Gao, *Phys. Rev. Lett.* 85, 622 (2000).
110. J. W. Mintmire, B. I. Dunlap, and C. T. White, *Phys. Rev. Lett.* 68, 631 (1992).
111. R. Saito, M. Fujita, G. Dresselhaus, and M. S. Dresselhaus, *Appl. Phys. Lett.* 60, 2204 (1992).
112. N. Hamada, S. Sawada, and A. Oshiyama, *Phys. Rev. Lett.* 68, 1579 (1992).
113. M. Schluter, M. Lannoo, M. Needels, G. A. Baraff, and D. Tománek, *Phys. Rev. Lett.* 68, 526 (1992).
114. N. Park, M. Yoon, S. Berber, J. Ihm, E. Osawa, and D. Tomanek, *Phys. Rev. Lett.* 91, 237204 (2003).
115. A. L. Mackay and H. Terrones, *Nature* 352, 762 (1991).
116. D. Vanderbilt and J. Tersoff, *Phys. Rev. Lett.* 68, 511 (1991).
117. S. T. Hyde, Crystalline frame works as hyperbolic films, in "Defects and Processes in the Solid State" (J. D. FitzGerald and J. N. Boland, Eds.). Elsevier, Amsterdam, 1993.
118. S. T. Hyde and M. O'Keeffe, *Philos. Trans. R. Soc. London Ser. A* 354, 1999 (1996).
119. S. Bandow, F. Kokai, K. Takahashi, M. Yudasaka, and S. Iijima, *Appl. Phys. A* 73, 281 (2001).
120. A. Krishnan, E. Dujardin, M. M. Treacy, J. Hugdahl, S. Lynum, and T. W. Ebbesen, *Nature* 388, 451 (1997).
121. J. Tersoff, *Phys. Rev. B* 46, 15546 (1992-I).
122. R. F. Curl, in "Buckminsterfullerenes" (E. W. Billups and M. A. Ciufolini, Eds.). VCH, New York, 1993.
123. P. W. Fowler and D. E. Manolopoulos, "An Atlas of Fullerenes." Clarendon Press, Oxford, 1995.
124. T. W. Ebbesen, *Phys. Today* 49, 26 (1996).
125. M. M. Treacy, T. W. Ebbesen, and J. M. Gibson, *Nature* 381, 678 (1996).
126. M. Gomberg, *J. Am. Chem. Soc.* 22, 757 (1900).
127. A. V. Zenkevich, M. A. Pushkin, V. N. Tronin, V. N. Nevolin, G. A. Maximov, D. O. Efimov, and E. Lagsgaard, *Phys. Rev. B* 65, 073406 (2002).
128. A. Perez, P. Melinon, V. Dupuis, P. Jensen, B. Prevel, J. Tuaillon, L. Bardotti, C. Martet, M. Treilleux, M. Broyer, M. Pellarin, J. L. Vaille, B. Palpant, and J. Lerme, *J. Phys. D: Appl. Phys.* 30, 709 (1997).
129. H. Haberland, Z. Insepov, and M. Moseler, *Phys. Rev. B* 51, 11061 (1995).

130. W. Harbich, S. Fedrigo, and J. Buttet, *Chem. Phys. Lett.* 195, 613 (1992).
131. W. Harbich, S. Fedrigo, J. Buttet, and D. M. Lindsay, *J. Chem. Phys.* 96, 8104 (1992).
132. Y. Qiang, R. F. Sabiryanov, S. S. Jaswal, Y. Liu, H. Haberland, and D. J. Sellmyer, *Phys. Rev. B* 66, 064404 (2002).
133. S. Vijayalakshmi, A. Lan, Z. Iqbal, and H. Grebel, *J. Appl. Phys.* 92, 2490 (2002).
134. G. Deutscher, A. M. Goldman, and H. Micklitz, *Phys. Rev. B* 31, 1679 (1985).
135. C. Roland, V. Meunier, B. Larade, and Hong Guo, *Phys. Rev. B* 66, 035332 (2002).
136. L. N. Dinh, S. E. Hayes, A. E. Wynne, M. A. Wall, C. K. Saw, B. C. Stuart, M. Balooch, A. K. Paravastu, and J. A. Reimer, *J. Mater. Sci.* 37, 3953 (2002).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanostructures Within Porous Materials

Y. Kumzerov, S. Vakhrushev

Ioffe Physico-Technical Institute, St. Petersburg, Russia

CONTENTS

1. INTRODUCTION

The experimental implementation of new effects in the physics of nanostructures relies upon our ability to create new types of structures and devices. Our understanding of material processing in the pursuit of ultrasmall structures is steadily advancing. Epitaxial growth and lateral microstructuring techniques have made it possible to create low-dimensional electronic systems with quantum-confined structures, that is, quantum wells, quantum wires, and quantum dots. Nanostructures can also be obtained by confining a solid or a liquid within nanometer-sized pores of various porous materials. There are a lot of materials that incorporate into their structure systems of voids (nanometer pores), and, what is more, we may say that virtually all materials have some porous structure if they are not single crystals. When we fill such a porous material with some substances, we may prepare some types of nanostructures. Systems with nanoparticles embedded in a porous matrix via chemical coating have received some attention in recent years. In this case, the confined liquid materials penetrate into the pores by means of wetting processes. In the case of nonwetting there is a possibility of using a mechanical coating of inner surfaces of porous materials by forcing a nonwetting liquid into the pores by external pressure. The material coating the inner surfaces of the pores in different porous matrices may be of interest as a physical object with characteristic nanometer sizes, that is, some new types of nanostructures. We emphasize some advantages of such nanostructures:

1. It is possible to prepare nanostructures with a large range of characteristic sizes (from approximately 1 nm to approximately 200 nm).
2. It is possible to prepare nanostructures with various geometries (two-dimensional (2D) film-like structures, 1D wire-like structures, 0D small particle-like structures).
3. It is possible to prepare nanostructures from various substances (metals, semiconductors, insulators, etc.).
4. It is possible to prepare nanostructures with a very large amount of nanomaterial (up to several cubic centimeters). This situation permits the use of some experimental methods (e.g., neutron scattering, heat capacity measurements, etc.) that require a large amount of nanostructures.

It seems that such nanoobjects are similar to the well-known nanocomposite materials, used very widely nowadays, and in some situations it is not so simple to distinguish them. We mean that nanocomposites are rather 3D objects demonstrating new special properties owing to some interaction or combination of properties intrinsic to the nanoporous matrix and to a substance introduced into nanopores. In the case of nanostructures within nanoporous matrices, we are mainly interested in the properties of a substance in nanovoids and take into account the porous matrix virtually only as some external support for such a nanostructure.

Recently there have been a lot of investigations (more than 500 in the past 10 years) that can be considered to be concerned with nanostructures of this kind. At the same time, there are no reviews describing various aspects of the physics of nanostructures within porous materials. The objective of this chapter is to fill the gap in the existing literature.

This chapter is organized as follows. Examples of nanoporous matrices are presented in Section 2 with some descriptions of their porous structures, including characteristic sizes of nanopores, distributions of characteristic sizes of nanopores, relative volumes of nanopores, etc. Section 3

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (811–849)

describes some methods for the preparation of nanostructures within nanoporous matrices, including mechanical methods for which such a characteristic as wetting is very important, and some chemical methods, electrochemical methods, etc. Section 4 is devoted to various nanostructures based on nanoporous matrices and to physical properties of such nanostructures and possible applications.

2. POROUS STRUCTURES

2.1. Porosity

The "porosity" of a solid material means the existence of some free volume inside its structure that is not occupied by structural elements of this material. Such a porous volume can be divided ambiguously into elementary porous structures—pores, which may differ in their characteristic sizes, types of shape, and types of connections. They may be organized as an open porous system as to be isolated. The possibility of having different chemical compositions, different structural units, and different origins of porosity leads to a very large diversity of porous structures. The porous structures can be conventionally divided, according to their genesis, into two large groups: systems based on a composition of some solid elements with porous space situated between these elements and those based on extraction of some part of structural elements due to decomposition, dissolution, or leaching from a bulk nonporous material. Examples of the first group are tissues, paper, granules of sorbents and catalysts, porous ceramics, crystals of zeolites and chrysotile asbestos, etc. Examples of the second group are dry hydrates of metals, carbonized and activated coals, porous glasses, porous membranes, etc. The general characterization of a porous system is determined by the characteristic sizes of its pores and by geometrical factors. There are porous systems with strongly regular alternation of identical pores connected through identical channels and those in which the characteristic sizes of pores, their arrangement, and connections are stochastic.

The quantitative characterization of porous materials can be based on:

1. The porosity of a porous system ϕ, which can be defined as the relative volume of the porous space, that is, the volume of the porous space, V_0, divided by the total volume V: $\phi = V_0/V$. The relative volume of the nonporous skeleton is in this case $1 - \phi$. Occasionally, such a parameter as $\phi/(1 - \phi)$ is used to estimate the porosity. For the existing porous materials the value of ϕ may vary between virtually 0 to approximately 0.9.
2. The characteristic sizes of pores, which reflect the typical effective sizes of cross sections of the empty porous space and are the most important parameters controlling the main processes in a porous medium. To have a simple parameter, the complicated form of the porous space can be reduced to the simplest geometrical figures with some equivalent parameters (spheres or cylinders with characteristic diameters, etc.). According to the regulations of the International Union of Pure and Applied Chemistry (UIPAC) porous materials are named according to their characteristic pore sizes: submicroporous (diameter less than 0.4 nm), microporous (diameter in the range 0.4–2 nm), mesoporous (diameter in the range 2–50 nm), and macroporous (diameter exceeding 50 nm).
3. The size distribution of pores. This characteristic is often stochastic and can be determined through such a function as the density of diameter distribution $f(D)$, where $f(D)\,dD$ is the probability of finding pores with diameters in the range of D to $D + dD$ in the unit volume of a porous material. Naturally, $\int f(D)\,dD = 1$. There is a more practically interesting function $V(D)$—the distribution of the porous volume over diameters. $V(D) = \int (D^2/D^2) f(D)\,dD$, where $D^2 = \int D^2 f(D)\,dD$.
4. The inner surface of a porous material, which corresponds to the area of the interface between the porous space and the skeleton of the matrix. It can be defined as the relative surface S, that is, the internal surface area per unit mass or unit volume of a porous material (measured in m^2/g or m^2/cm^3). For different porous materials, S may vary between about 10^{-2} and about 10^3 m^2/g.

Naturally, real porous objects mainly have irregular and stochastic systems of nonidentical voids with parameters that are very difficult to calculate. In this case, a usual way to describe them is associated with attempts to reduce the porous system to various geometrical models. A simple variant is to reduce porous materials to systems of spheres, modeling the skeleton of a porous matrix, in which the empty space between these spheres is modeling the real porous space. In the simplest case of identical spheres with regular packing, the needed parameters are only the sphere radius R and the coordination number n, characterizing the density of packing. In this case, the main parameters for a porous medium can be found as [1]

$$\phi = \frac{2.62}{n}; \qquad S = 3(1 - \phi)R^{-1} \tag{1}$$

Unfortunately, the simplest geometrical models cannot describe the topologically complex space of real, irregular porous matrices. A stronger description of porous systems of this kind was given in [2], where it the porosity itself was regarded as a fundamental random variable, because the value of porosity ϕ or the pore size distribution is not sufficient to characterize the complex pore-space geometry. The characterization is based on analyzing the local geometry on a mesoscopic scale, and the local geometry around a lattice point $\mathbf{R}$ is defined as the intersection of the pore space and the primitive cell at $\mathbf{R}$. To define the local porosity, the characteristic function of an arbitrary set A is introduced as $\chi_A(\mathbf{r}) = 0$ if $\mathbf{r}$ lies outside the set A and 1 if $\mathbf{r}$ lies inside the set A.

The local porosity $\phi(\mathbf{R}, L)$ at the lattice position $\mathbf{R}$ is defined for length scale L as

$$\phi(\mathbf{R}, L) = \rho \int \chi_{\mathrm{MC}}(\mathbf{r}; \mathbf{R}, L)\chi_{\mathrm{PS}}(\mathbf{r})\,d\mathbf{r} \tag{2}$$

where ρ is the density of Bravais lattice points; $\chi_{\mathrm{MC}}(\mathbf{r}; \mathbf{R}, L)$ is the characteristic function of the measurement cell at $\mathbf{R}$, having size L; and $\chi_{\mathrm{PS}}(\mathbf{r})$ is the characteristic function of

the porous space. The integration is done over the porous medium. To define the local porosity distribution function, let $\mu(\phi, \mathbf{R}; L)$ measure the probability of finding the local porosity ϕ in the range from ϕ to $\phi + d\phi$ in a cell of linear dimension L at a point $\mathbf{R}$. If the porous medium is homogeneous, μ must be independent of $\mathbf{R}$, and $\mu(\phi, \mathbf{R}; L) = \mu(\phi; L)$ can be named the local porosity density at scale L. The bulk porosity ϕ^* can be determined as an integral over the entire volume or as an average over a statistical ensemble of measurement cells, and, assuming "ergodicity," we have $\phi = \phi(\mathbf{R}, L \to \infty) = \int_0^1 \phi_\mu(\phi; L)\, d\phi$, independent of $\mathbf{R}$ and L.

Higher-order distribution functions $\mu_n(\phi_1, \mathbf{R}_1; \phi_2, \mathbf{R}_2; ; \phi_n, \mathbf{R}_n; L)$ measure the probability density of finding ϕ_1 in the cell at $\mathbf{R}_1$, ϕ_2 in the cell at $\mathbf{R}_2$, etc. at scale L. A set of such functions gives full information about the statistical properties of a porous medium at scale L. The local porosity distribution $\mu(\phi; L)$ depends on L. There is an intermediate length scale ξ, the porosity correlation length, at which the local geometries are relatively simple, but, on the other hand, the single-cell distribution function has sufficient nontrivial geometric content to be a good approximation. More precisely, ξ is determined by a two-cell distribution function, $\mu_2(\phi_1, \mathbf{R}_1; \phi_2, \mathbf{R}_2; L)$, which depends only on the distance R in the isotropic case, that is, $\mu_2 = \mu_2(\phi_1, \phi_2; R; L)$. Namely, $\xi^2 = \{\int R^2 C(R, 0) d^3R\} / \{\int C(R, 0) d^3R\}$, where $C(R, L)$ is the porosity autocorrelation function at scale L:

$$C(R,L) = \left\{ \int_0^1 \int_0^1 (\phi_1 - \phi^*)(\phi_2 - \phi^*) \mu_2(\phi_1, \phi_2; R; L) d\phi_1 d\phi_2 \right\} \Bigg/ \left\{ \int_0^1 (\phi - \phi^*)^{2\mu} (\phi; R; L) d\phi \right\}. \quad (3)$$

The local porosity distribution can be defined as $\mu(\phi) = \mu(\phi; \xi)$, the single-cell porosity at scale ξ. The important aspect of $\mu(\phi) = \mu(\phi; \xi)$ is that it is measurable with modern image-processing equipment. The procedure for observing $\mu(\phi)$ in homogeneous and isotropic media measures this function on photographic images of thin 2D sections across the porous space. The local porosity distribution contains a good deal of geometrical information about the pore-space geometry, and it is optimal in the sense that it contains the maximum amount of information based purely on the porosity concept. The second geometric property to characterize local geometries is whether or not the pore space percolates. Let $\lambda(\phi)$ denote the fraction of percolating cells with local porosity ϕ. $\lambda(\phi)$ can be called the local percolation probability. The two functions $\mu(\phi)$ and $\lambda(\phi)$ constitute only a partial and approximate geometric characterization of the pore space, but they have rich geometrical content.

2.2. Characterization of Porous Media

For the pore sizes of the nanometer scale there is no unique method yielding full information about the porous structure. We need to use various methods to extract the whole body of information of interest [1, 3, 4]. The most essential of these are listed in Table 1.

2.2.1. Electron Microscopy

This is the well-known method for the study of porous structures (see [3, 4]), frequently used as an additional method because of the laborious procedure of sample preparation, especially in the case of transmission electron microscopy. This method gives in its simplest form direct information about the sizes and shapes of pores, and about their connections, but only in some local positions. For more complete information, several measurements are needed for varying and averaging the positions on a sample, as are the feature-counting measurements on electron micrographs. As an example of a study of this kind, mention can be made of the fact that the fractal geometry was confirmed for the porous space in sandstone with the use of scanning electron microscopy (SEM) with manual feature-counting in SEM images [5]. For the same purposes, an automated technique

Table 1. Information about methods for porous media characterization.

Method	Information that can be obtained	Minimum characteristic size
1. Electron microscopy	Amount, volume, and distribution of pores. Internal surface area. The possibility to determine the closed pores. There is only local information.	≈0.5 nm
2. Methods based on porous structure filling with wetting materials		
a. Pycnometry	Total porosity	≈0.5 nm
b. Monolayer and multilayer adsorption	Internal surface area	≈0.5 nm
c. Capillary condensation	Volume and distribution of pores, internal surface area	≈1 nm
3. Methods based on porous structure filling with nonwetting materials (mercury porosimetry)	Volume and distribution of pores, internal surface area	≈1 nm
4. Small-angle scattering of X-rays and neutrons	Pore sizes, internal surface area. The possibility to determine the closed pores. Nondestructive method	≈0.5 nm
5. Methods based on NMR		
a. Relaxation time measurement for confined liquids	Volume and distribution of pores, internal surface area	≈2 nm
b. Pulsed field gradient NMR	Pore sizes, internal surface area	

has also been developed for making SEM measurements [6] with image digitizing, filtering, counting of geometrical features as a function of feature sizes, and fitting feature histograms.

2.2.2. Methods Based on Filling of Porous Structures with Wetting Materials

Pycnometry This method for studying porous structures [3, 4] is based on a comparison of the densities of porous materials with empty and filled porous space. The pores in this method may be very easily filled with a liquid that wets the internal surface of the porous space (in the simplest case this may be water) and thereby can penetrate all open pores. The difference between the densities of a sample with empty pores and the same sample with filled pores gives the total porous volume for this sample if the density of the wetting liquid is known. Thus this method permits one to determine the total porosity of porous materials.

Monolayer and Multilayer Adsorption A study of adsorption of various gases by porous materials is a source of important information about the internal surface of a porous space (see [3, 4, 7–12]). This process can be described by isotherms of adsorption, that is, by the dependences of the volume of adsorbed substance on the relative pressure at a given temperature. The substance most frequently used for adsorption in nanoporous materials is nitrogen at 77 K.

A theoretical description of monolayer adsorption is based on a model developed by Langmuir [13], in which the isotherm looks like

$$V/V_{\mathrm{m}} = (Bp/(1 + Bp)) \tag{4}$$

Here V is the volume of condensed sorbate, V_{m} is the maximum volume corresponding to the inner surface covered with a monolayer of sorbate, $p = P/P_0$ is the relative pressure, P_0 is the saturated vapor pressure, and B is a constant. From experimental data, $V = V(p)$, it is very easy to determine first V_{m} and then S, the internal surface area of the porous space, if we know the characteristic sizes of molecules for the sorbate used. Brunauer et al. [14] developed this model for multiple-layer adsorption. The isotherm of adsorption corresponds in this case to the equation (the so-called BET equation)

$$V/V_{\mathrm{m}} = \{Cp/(1 - p)[1 + (C - 1)p]\}, \tag{5}$$

where $C \approx \exp(Q/RT)$, Q is the heat of adsorption, R is the gas constant, and T is temperature. This equation, when written in the form $p/V(1-p) = 1/V_{\mathrm{m}}C + p(C-1)/CV_{\mathrm{m}}$, is very convenient for the approximation of experimental data, because the dependence of $p/V(1 - p)$ on p, which can be obtained in experiment, is a straight line, which permits one to determine the coefficients $1/V_{\mathrm{m}}C$ and $(C - 1)/CV_{\mathrm{m}}$ and then to calculate V_{m} and C. It was established that experimental data are frequently in agreement with the BET equation, especially at relative pressures p in the range from approximately 0.05 to approximately 0.35.

Capillary Condensation When $p = P/P_0$ is more than approximately 0.75, the pore space is filled by sorbate not through multilayer adsorption but through capillary condensation. This results from a decrease in the saturated vapor pressure over the curved surface of the wetting liquid in a capillary in comparison with the saturated vapor pressure over the planar surface of the same liquid, which leads to condensation of wetting liquids in the pores at $p < 1$. The condensed liquid volume can be calculated with the Kelvin equation (see [8, 9])

$$\ln(P/P_0) = -4\sigma V \cos\theta/DRT \tag{6}$$

where σ is the surface tension of the liquid, V is the molar volume of condensed liquid, θ is the wetting angle (at full wetting $\theta = 0$ and $\cos\theta = 1$), D is the pore diameter, R is the gas constant, and T is the temperature. The experimental dependence $V = V(P)$ can be used to calculate, with the Kelvin equation, the dependence of the characteristic porous volume on pore diameter. A more exact description of capillary condensation needs to take into account the thickness of adsorption layers and the dependence of the surface tension on the curvature of the liquid meniscus (see [9]). The adsorption and desorption processes also exhibit a hysteresis depending on the geometry of the porous space. Commonly, practical calculations use the simplest form of the Kelvin equation and the desorption branch of the capillary condensation hysteresis. Some examples of such a study are the following: [15] for xenon adsorption on silica gel; [16, 17] for adsorption of Ar, N_2, O_2, C_2H_4, and CO_2 on model mesoporous adsorbents MCM-41; and [18] for the adsorption of xenon on Vycor porous glass. Some models have been proposed for describing the adsorption hysteresis on the basis of Monte Carlo [19] and molecular dynamic computer simulation [19, 20], the density functional approach [21], the lattice gas model [22], and percolation theory [23].

2.2.3. Mercury Porosimetry

This method is based on the possibility of filling porous materials under high pressure with liquids that do not wet the surface of the porous space (see, e.g., [3, 4, 24]). Owing to the property of mercury that it will not wet most porous materials, it has frequently been used for these purposes, and this method has been named "mercury porosimetry" (some other materials can also be used, see, e.g., [25], where the Wood alloy was employed). The pressure needed for filling is the well-known Laplace pressure. In the simplest case (for full nonwetting, when the wetting angle θ is 180°) it can be written as $P = 4\sigma/D$ (σ, surface tension; D, pore diameter). At this pressure the entire porous volume of pores with diameters exceeding D is filled with nonwetting liquid. From the experimental dependences of filled porous volume on external pressure it is possible to calculate the distribution of porous volume over diameters. $P = 4\sigma/D$ can be written as $4F/\pi D^2 = 4\sigma/D$ (F strength). For a cylindrical pore with length L this expression can be rewritten as $4FL/\pi D^2 = 4\sigma L/D$ or $FL = \sigma(\pi DL) \approx \sigma S$, where FL is

the mechanical work done by the piston pressing liquid into the pore and S is the total surface area of the cylinder. In this form the above equation reflects the process of mechanical energy transformation during the process of filling into the surface energy σS of dispersed liquid and permits one to determine the total surface area of the porous space. When there is some distribution of pore diameters, the surface area of the pore space can be determined from the experimental dependence of filled volume on the external pressure $V(P)$

$$S = (-1/\sigma)\int_0^{V\max} P\,dV \tag{7}$$

2.2.4. Small-Angle Scattering of X-Rays and Neutrons

X-ray and neutron scattering is a powerful technique for studying crystalline materials. For a given wavelength λ, the scattering angles φ for a lattice constant d are determined by the Bragg condition $n\lambda = 2d\sin\varphi$, where n is an integer. In porous materials, the characteristic lengths $L \gg d$, and the scattering features are observed at small angles. For example, for $\lambda \approx 0.15$ nm and 2 nm $< L <$ 100 nm the typical angles φ are in the range from 5° to 7°. A theory has been developed for describing the X-ray and neutron scattering at small angles (see [26, 27]), and such a study is very informative regarding porous structures at the nanometer scale (see [3, 4, 28–30]). According to a theory for scatterers bounded by a smooth surface, the limiting form of the scattered intensity must be proportional to q^{-4}, where $q = 4\pi\lambda^{-1}\sin(\varphi/2)$. Some deviations from this dependence have been observed for various porous materials (sedimentary rocks, lignite coals, silica aerogels, porous glasses) and attributed to the existence of fractals (for fractals, see [31]) in the porous structure of these materials (see [32–38]). A difference between volume and surface fractals was established. In the former case, the scattering intensity $I(q) \sim q^{-D_v}$, where D_v is the fractal dimension relating the mass M to the characteristic length L as $M(L) \sim L^{D_v}$ for a sphere of radius L. In the latter case, the scattering intensity $I(q) \sim q^{D_s-6}$, where D_s is the surface fractal dimension which relates the surface area S to length as $S \sim L^{D_s}$ for a sphere of radius L. The conventional dependence $I(q) \sim q^{-4}$ corresponds to $D_s = 2$, that is, to a 2D geometry of porous surface space.

2.2.5. Methods Based on Nuclear Magnetic Resonance

Measurement of Relaxation Time for Confined Liquids

Relaxation time measurements can probe the local surface-to-volume ratio for a porous space and thus can determine the average pore sizes and pore distributions [39–41] (see, also [42]). In this method, the relaxation of magnetization with time is measured after the application of an appropriate pulse sequence to a porous sample filled with some suitable liquid. An increased relaxation rate (either transverse or longitudinal) caused by the interaction of the solid surface with fluid molecules near the surface is assumed for molecules within a certain distance λ from the surface. For the liquid within a single pore, there is a single relaxation time $1/T = 1/T_B + (s/v)(\lambda/T_S)$, where $1/T$ is the longitudinal or transverse relaxation rate; $1/T_B$ and $1/T_S$ are, respectively, the corresponding relaxation rates in the bulk or within λ from the surface; s and v are the surface area and volume of the pore. For the entire sample the contribution of relaxation rates from different pores to magnetization decay can be written in the continuum form,

$$A(t) = \int_0^\infty P(D)\exp[-t(1/T_B + (\lambda/T_S)(m/D))]\,dD \tag{8}$$

where $A(t)$ is reduced magnetization which depends on the NMR signal $E(t)$: $A(t) = (1/2)[1 - E(t)/E(\infty)]$ for a longitudinal inversion recovery sequence and $A(t) = E(t)/E(0)$ for a transverse pulse sequence ($E(\infty)$ is the echo signal upon full recovery and $E(0)$ is that at zero delay time), $P(D)$ is the pore size distribution function, D is the pore diameter, $s/v = m/D$ (the constant m depends on the pore geometry). Inversion of the experimental magnetization data for practical purposes may permit calculation of the function $P(D)$. In [41], the experimental $P(D)$ data determined in relaxation time measurements with various porous glasses were compared with the experimental results obtained with measurements by the adsorption method and mercury porosimetry.

Pulsed-Field-Gradient NMR Pulsed-field-gradient (PFG) nuclear magnetic resonance (NMR) is a versatile tool for studying the transport of fluids in porous media [43–47]. The PFG NMR technique can reveal structural properties of the porous medium by measuring the molecular transport, that is, by monitoring the restricted self-diffusion of the molecules. NMR, combined with an additional space- and time-dependent magnetic field, so-called pulsed field gradients (PFGs), offers an opportunity to generate labeled molecules and to detect their mean displacements related to any transport process (the labeling is due to the Larmor precession ω of nuclear spins associated with the molecules). The Larmor frequency $\omega(z, t)$ can be written as $\omega(z, t) = -\gamma(B_0 + g(t)z)$, where B_0 is time-independent uniform magnetic field parallel to the z coordinate, $g(t)$ is an additional time-dependent magnetic field with a field gradient oriented parallel to the z direction, and γ is the gyromagnetic ratio of the nuclei under investigation. In such a magnetic field, the complex magnetization density $m(\mathbf{r}, t)$ gets an additional z dependence,

$$m'(\mathbf{r}, t) = m(\mathbf{r}, t)\exp\left(-i\gamma z\int_0^t g(t')\,dt'\right) \tag{9}$$

If the magnetic field gradient $g(t)$ is turned off after a time δ, the magnetization density is found to be embossed with a z-dependent factor

$$m'(\mathbf{r}, t) = m(\mathbf{r}, t)\exp(-i\gamma g\delta z) \tag{10}$$

This phase factor represents the labeling of the molecules observed in PFG NMR diffusion studies. A length scale $l_{PFG} = (2k)^{-1}$ ($k = \gamma\delta g$) can be introduced, over which a significant change in the phase factor and, hence, the labeling of the diffusing molecules are achieved. The self-diffusion process of the molecules observed smears out the

originally present spatial distribution of phases. Measurement of this phase distribution after an observation time Δ allows evaluation of the self-diffusion of the molecules. This process is caused by a second magnetic field gradient pulse after a time Δ, which has to refocus the special phase distribution introduced by the first pulse. Owing to the displacements of the molecules, caused by self-diffusion, the refocusing is not complete. The NMR signal observed, which follows from integration of the magnetization density over the sample volume, is reduced by the factor

$$\psi(k, \Delta) = \int P(z - z, \Delta) \exp[ik(z - z')]d(z - z')$$

in comparison with the signal measured without a pulsed field gradient ($P(z - z', \Delta)$ denotes the probability density that a molecule is displaced by $z - z'$ within a time interval Δ). $P(z - z', \Delta)$ is given by the inverse Fourier transform of the NMR signal attenuation [48] and contains information about the geometrical features of the pore space, which affect the molecular motion of the pore fluids [49–53]. The power of PFG NMR as a tool for the investigation of porous systems is based on the fact that l_{PFG} and k can easily be controlled by choosing the appropriate values for the pulsed field gradient intensity g and width δ. l_{PFG} values of less then 0.1 μm (required, e.g., for ^{1}H NMR with $\delta \geq 1$ ms and $g \geq 20$ T/m) can be achieved with the current state-of-the-art PFG NMR spectrometers [54].

2.3. Porous Matrices

2.3.1. Matrices with Irregular Porous Structure

Porous Glass A porous silica glass can be defined as a bicontinuous random structure of two interpenetrating percolating phases, namely the solid and the pore networks. The pores in the glasses are connected to each other, and the pore size distribution is narrow (see Fig. 1). Porous glass is available in standard forms such as tubing, rods, and sheets. It is mechanically hard and strong, nondusting, nonflaking, and chemically inert (for its physical and chemical properties, see [55–62]). The preparation methods and properties of various types of porous glasses are summarized in [63] (see also [64–77]). Porous glass is often used for the filtration and separation of compounds. The open-cell network allows permeability on a selective basis—the species must be smaller than the microscopic pores to pass through a porous glass. A porous glass matrix can also be used in nuclear waste storage [78, 79] and optical applications [80]. The present status and future potential for porous glass applications are reviewed in [81]. The homogeneous pore diameters can be controlled to be, on the average, between 40 and 500 Å. A typical example of porous silica glasses is Vycor glass (Corning vycor brand 7930 [82]). Its standard chemical composition is as follows: 96% SiO_2, 3% B_2O_3, 0.40/a Na_2O, $R_2O_3 \pm RO_2 < 1\%$ (R = mostly Al_2O_3 and ZrO_2). Its physical characteristics are given in Table 2.

For Vycor porous structure, see also [29, 36, 37, 75, 83–89]. Such a structure is obtained as a result of spinodal decomposition of the following two phases: SiO_2 and $B_2O_3 + Na_2O$. Vycor porous glass is prepared in three steps [90]. A melt composed of 75% SiO_2, 20% B_2O_3, and 5% Na_2O is quenched near (but below) its consolute temperature, inside the immiscibility gap. During the heat treatment, a slow liquid-liquid diffusion occurs, resulting in separation into two phases, rich in B_2O_3 alkali oxide and in SiO_2. After the glass is heat-treated and annealed, the B_2O_3-rich phase is removed by leaching at about 100 °C with acid, leaving an almost pure SiO_2 skeleton. The resulting structure of the interconnected matrix has been interpreted in different ways. First the pore network was suggested to be a mass fractal, associated with 3D percolation clusters [91]. Next, Vycor glass was analyzed [85, 92] with the theory developed for the early stages of spinodal decomposition [93]. In this model, the phase separation occurs by growth of unstable long-wavelength concentration fluctuations. The acid leaching removes the regions "poor" in SiO_2, creating an interconnected solid skeleton. The third proposed separation process is suggested to be driven by nucleation and growth of an almost pure SiO_2 phase [37, 94, 95]. The Vycor glass must have a structure made up of a randomly packed assembly of spheroidal particles forming a connected and homogeneous solid network. The geometrical nature of the interfacial microstructure of the pore network has also been controversial. There have been discussions in the literature, regarding both the volume and surface fractal structures in Vycor glass [36, 37, 83–86, 91, 96, 97]. Most of the information

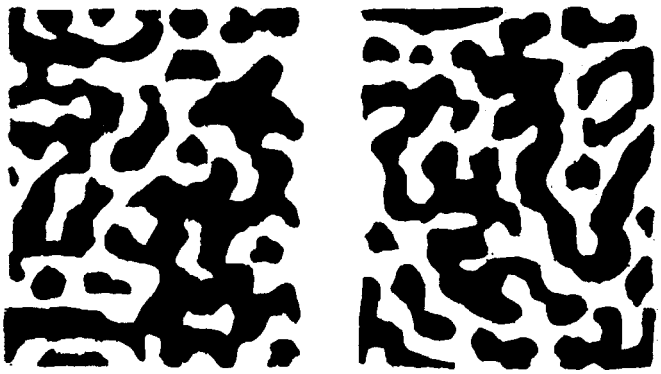


Figure 1. Model of porous structure for porous glass.

Table 2. Main physical characteristics of porous glasses.

Parameter	Value
Approx. specific gravity (dry)	1.5
Void space	28% of vol.
Internal surface area	250 M^2/g
Avg. pore diameter (standard)	40 Å
Appearance	Opalescent
Avg. modulus of rupture of abraded "A" rods, 25 °C	6000 psi
Young's elastic modulus, 25 °C	2.5×10^6 psi
Loss tangent at 25 °C, 100 Hz	0.007
Dielectric Constant at 25 °C, 100 Hz	3.1

Note: Loss tangent and dielectric constant are affected by water in porous glass.

related to the pore morphology has been obtained by means of transmission electron microscopy, adsorption methods, and small-angle scattering techniques. A study using a small-angle scattering technique [36] interpreted the power-law scattering at large q ($\sim q^{-4}$) to be a signature of scattering from a smooth interface [26]. However, some results [37] show an asymptotic behavior at large q, with power-law scattering of the form $q^{-3.7\pm0.1}$. This noninteger exponent may be due to scattering from a rough pore interface.

Xerogels and Aerogels Some other materials with an irregular porous structure are xerogels and aerogels. These materials are based on a gel that is on a suspension or polymer solution that behaves as an elastic solid or semisolid, rather than a liquid. A dried-out gel is termed a xerogel. A xerogel has an open porous structure at the nanometer scale. The porosity of this material is due to a random colloid aggregation process in the solution precursor. The classical models of their porosity-dependent properties are based on highly simplified geometrical structures like packed spheres or bottlenecked bubbles [7]. An aerogel is a special kind of xerogel in which the dried-out gel retains most of the original open structure. The most frequently used xerogels and aerogels are those based on SiO_2: silica gels and silica aerogels (for their production, see [98]). Their porous structure has been studied by various methods: gas/vapor adsorption (see [15]); mercury porosimetry (see [24], but this technique is generally not effective with aerogels); X-ray, neutron, and visible light scattering methods (see [99–108]); and numerical simulations [109–111]. Pores in various materials are either open or closed, depending on whether the pore walls are solid or porous themselves. A macroscopic example of an open-pore material is a common sponge. In a closed-pore material, gases or liquids cannot enter a pore without breaking it. This is not the case with an open-pore structure. In this instance, gases or liquids can flow from pore to pore with limited restriction, and eventually through the entire material. It is this property that makes aerogels effective materials for gas-phase catalysts, microfiltration membranes, and substances for chemical vapor infiltration. Physical properties of typical silica aerogels are listed in Table 3.

Table 3. Physical properties of silica aerogels.

Apparent density	0.003–0.35 g/cm^3	Most common density is about 0.1 g/cm^3.
Internal surface area	600–1000 m^2/g	As determined by nitrogen adsorption/desorption.
% solids	0.13–15%	Typically 5% (95% free space).
Mean pore diameter	≈20 nm	As determined by nitrogen adsorption/desorption.
Primary particle diameter	2–5 nm	Determined by electron microscopy.
Coefficient of thermal expansion	$(2–4)*10^{-6}$	Determined with ultrasonic methods.
Dielectric constant	≈1.1	For density 0.1 g/cm^3. Very low for solid material.
Sound velocity through the medium	100 m/s	For density 0.07 g/cm^3. One of the lowest velocities for a solid material.

Nanoporous Particle Track-Etched Membranes In 1962, Price and Walker [112, 113] made an interesting observation that the tracks in a material damaged by the passage of heavy charged particles through insulators can be revealed by chemical etching to form pores. The tracks that traverse the specimen are seen to be quite uniform and about 700 Å in diameter. Bean et al. [114] described an electrical method for subsequent pore growth during etching and a theory describing the rate of pore enlargement. They used clear muscovite that was cleaved to thicknesses between 3 and 8 μm. Irradiation was accomplished from a sample of spontaneously fissionable ^{252}Cf. About 10^5 tracks/cm^2 were formed in 16 h. The process of pore growth was monitored by measuring conductance across a thin sample. In summary, this study demonstrated that pores in mica, with radii from 30 Å to many micrometers can be created and monitored. A similar technology has been developed for polymeric track and other membranes [115–121]. The technology consists of irradiation of thin polymer foils with accelerated heavy ions, sensitization, and chemical treatment, resulting in the formation of uniform micro- or nanopores in the polymeric matrix. The pores are straight cylindrical channels with a clearly defined diameter. The surface pore density is also precisely predetermined by irradiation conditions. Special etching procedures give pore channels as small as 15–50 nm. The pore shape can also be changed at will. At the Joint Institute of Nuclear Research (Dubna, Russia) [115, 116] the irradiation of polymer foils is carried out on a U-400 cyclotron. It furnishes an opportunity to perform a homogeneous treatment of large sample areas. Accelerated Kr and Xe ions with energies of 4–6 MeV/amu are available.

2.3.2. Matrices with Regular Porous Structure

Zeolites Zeolites are nanoporous crystalline solids with well-defined structures [122–126]. Generally, they contain silicon, aluminum, and oxygen in their skeleton. Many zeolites occur naturally as minerals and are extensively mined in many parts of the world. Others are synthetic and are made commercially for specific uses [126–129]. Natural and synthetic zeolites are typically powders with a characteristic size less than 10^{-2} cm. Because of their unique porous properties, zeolites are used in a variety of applications with a global market of several million tons per annum. Their major applications in the world are in petrochemical cracking, ion exchange (water softening and purification), and separation and removal of gases and solvents. Other applications are in agriculture, animal husbandry, and construction. Zeolites are often also referred to as molecular sieves. A defining feature of zeolites is that their frameworks are made up of 4-connected networks of atoms. One way of thinking about this is in terms of tetrahedra, with a silicon atom at the center and oxygen atoms at the corners. These tetrahedra can then be linked together by their corners to form a rich variety of structures. The framework structure may contain linked cages, cavities, or channels, which are roughly between 3 and 15 Å in diameter (see Fig. 2). The centers of zeolite cages (or channel axis) are arranged in space with crystalline order, which allows preparation of ensembles with size-calibrated clusters and/or atomic diameter wires with long-scale macroscopic ordering [130]. In all,

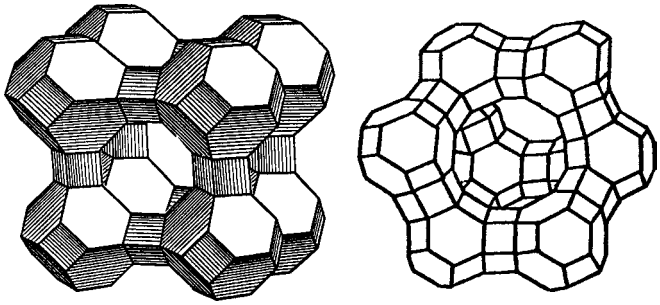


Figure 2. Zeolite cages for NaA (left) and NaX (right) matrices.

130 different framework structures are now known [131]. In addition to those having silicon or aluminum as the tetrahedral atom, other compositions have also been synthesized, including the growing category of nanoporous aluminophosphates known as ALPOs. The structural properties of some typical zeolites are listed in Table 4 [122, 130].

Mesoporous Molecular Sieves There are new kinds of nanoporous materials with regular systems of nanovoids, the so-called mesoporous molecular sieves, such as MCM-41 [132–137] and SBA-15 [138], which have many desirable properties for application as separation media (see also [139–145]). Members of this family of materials, designated MCM-41, were first observed in electron micrographs of products from hydrothermal reactions of aluminosilicate gels in the presence of quaternary ammonium surfactants. The material possesses a regular array of uniform, hexagonally shaped channels with narrow pore size distribution, the dimensions of whcih can be tailored in the range from 16 Å to 100 Å or more through the choice of surfactant, auxiliary chemicals, and reaction conditions. The pore walls are composed of amorphous silica, as indicated by X-ray diffraction measurements. MCM-41 has been used as a model system for sorption isotherms of various gases. The BET surface area is $\geq$1000 m^2/g with exceptionally high sorption capacities. The range of pore volumes for MCM-41 samples is 0.7–1.2 cm^3/g (porosity of up to 80%). The use of amphiphilic triblock copolymers to direct the organization of polymerizing silica species has resulted in the preparation of well-ordered hexagonal mesoporous silica structures, SBA-15, with uniform pore sizes up to approximately 30 nm. These materials are synthesized in acidic media to produce highly ordered, 2D hexagonal silica block-copolymer mesophases. Calcination at 500 °C gives porous structures with pore sizes from 4.6 nm to 30 nm, pore volume fractions up to 0.85, and silica wall thicknesses of 3.1 nm to 6.4 nm.

Artificial Opals Natural opal is known as a precious stone with beautiful colors due to its structure, which is really a natural diffraction grating for visible light [146–149]. It consists of identical silica spheres with a very narrow size distribution (for its structure, see also [150–153]). There are now artificial opals [147, 154, 155] that also consist of identical silica spheres with diameters (D) varying from 150 to 900 nm [156]. The size uniformity of these spheres allows their assembly in a close 3D lattice, usually with FCC symmetry (see Fig. 3). There are empty voids between neighboring spheres, which, in turn, form their own regular lattice. There are two types of interpenetrating voids in the opal lattice: 8-fold-coordinated large voids, each connected with eight 4-fold-coordinated small voids. The large void has the form of a truncated tetrahedron with four windows to four large voids. The diameters of spheres inscribed in the larger and smaller voids are $d_1 = 0.41D$ and $d_2 = 0.23D$, respectively. The diameter of a circle inscribed in the triangular window is $d_3 = 0.15D$. The density of voids in opal is typically 10^{14} cm^{-3}. The porosity of the ideal package of spheres is 26% of the whole volume. The main opal spheres consist of smaller spheres 30–40 nm in diameter, which in turn are composed of particles 5–10 nm in diameter. Thus, the total porosity of opals may approach 59%, constituted by the 26, 19, and 14% porosity corresponding to the lower hierarchy of size [157]. The actual porosity of the opal may be either above or below this value, depending on the synthesis history (sintering conditions, interstitial cement).

Chrysotile Asbestos Asbestos is a general term for a number of naturally occurring, fiber-rich minerals. The minerals are silicates, which means that the crystal structure is built up around SiO_4, frequently with a high concentration of magnesium, iron, or alkali metals [158, 159]. There are a number of health issues associated with exposure to asbestos [160, 161]. Chrysotile asbestos (chemical formula $Mg_3Si_2O_5(OH)_4$) is the only type of asbestos that can be

Table 4. Structural properties of typical zeolites.

Property	Zeolite NaX	Zeolite NaA	Zeolite NaM (mordenite)
Chemical formula	$Na_{86}[(AlO_2)_{86}(SiO_2)_{106}]32 \times 8H_2O$	$Na_{12}[(AlO_2)_{12}(SiO_2)_{12}]29H_2O$	$Na_{8.7}[(AlO_2)_{8.7}(SiO_2)_{39.3}]24H_2O$
Lattice of voids	Diamond-like with lattice constant 25.4 Å	Cubic with lattice constant 12.27 Å	Parallel channels with distance between their axes of 13–14Å
Diameter of voids	13 Å	11.4 Å	Diameters of channels 6.7–7 Å
Diameter of window between voids	7–8 Å	4.2 Å	
Relative volume of voids	42% of the total volume (there are some small voids with relative volume of 8%)	40% (there are some small voids with relative volume of 7%)	20%
Density with adsorbed water	1.91 g/cm^3	1.96 g/cm^3	2.13 g/cm^3
Density without water	1.41 g/cm^3	1.49 g/cm^3	

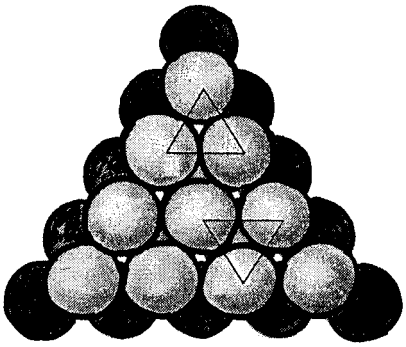


Figure 3. Packing of silica spheres in opal matrix.

decomposed by acids. It is a regular set of closely packed, parallel, ultrathin dielectric tubes with external diameters of about 20–25 nm. Its color may be white, gray, green, or yellowish. Currently chrysotile asbestos is the only type of asbestos mined on a large scale. The biggest deposit of asbestos mined now is located in the Urals (Russia) [162]. The mineral structure is described as a layer of partially hydrated MgO bound to a corresponding SiO_2. Since the MgO layer cannot adapt properly to the SiO_2 layer, being larger, the thin double layer rolls up into very thin tubes, with the magnesium oxide layer outermost. The fibers are therefore hollow cylindrical tubes with an outer diameter of 200–250 Å and an inner diameter of 20–50 Å [158, 159] that are convenient for nanowire preparation (see Fig. 4). The density of chrysotile is 2.4–2.6 g/cm^3. The surface area depends on the distribution of fibers and may vary within 4–50 m^2/g, which is more than the corresponding variation in other fibrous textile materials. The refractive index along the fibers is 1.531–1.541 [159, 163, 164]. The low-temperature magnetic susceptibility of Canadian chrysotile obeys the Curie law, and Rhodesian chrysotile has an antiferromagnetic Weiss constant of about 1 K. The apparent paramagnetism of chrysotile may be due to the presence of iron compounds, such as magnetite, as impurities [165]. The pH reaction of chrysotile corresponds to $Mg(OH)_2$ in water in that a 0.5% chrysotile suspension free of CO_2 has a pH of 10.33. The determination of the lattice parameters of chrysotile has been beset by unusual difficulties, as a result of the peculiarities of its diffraction pattern [166, 167].

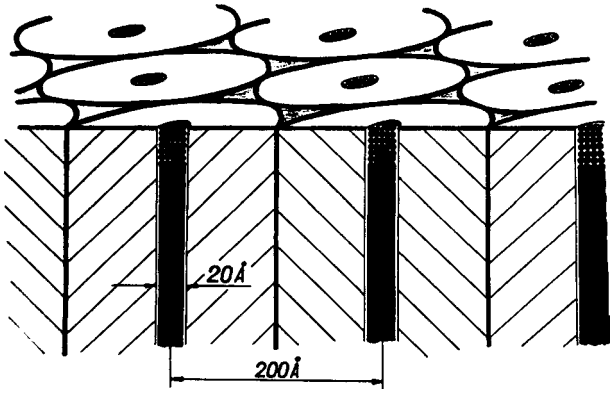


Figure 4. System of nanowires within the channels of chrysotile asbestos.

For example, no single crystals of the mineral are available, and the finest fiber specimen that can be examined gives a diffraction pattern with the full symmetry of a rotation photograph. Warren and Bragg [168] suggested a structure based on amphibole-like silicate chains. Warren and Herring [169] and Aruya [170] pointed out that the diffuse reflections of chrysotile are similar to cross-grating reflections and, therefore, assumed that the mineral has a disordered layered structure. Modern structural results were obtained by Whittaker [166] at a time when electron-microscopic results appeared to be well established, but before the theory of diffraction by cylindrically curved layers was worked out. The types of structure that can be formed by cylindrically curved layers have been classified, and their diffraction effects have been discussed systematically [171–175]. The typical specimen studied by Whittaker [176] was a fiber pencil (17 mm long and 0.14×0.07 mm in cross section) from Bell's mine, Thetford, Québec, Canada. All of the reflections from this specimen are in the positions expected for a monoclinic cylindrical lattice with parameters $a = 14.65 \pm 0.01$ Å, $b = 9.2$ Å, $c = 5.34 \pm 0.01$ Å, $\beta = 93°16' \pm 1'$. Some other types of chrysotile have also been studied by Whittaker [177–179] (for lattice parameters of chrysotile reported by different authors, see [180–184]. The porous structure of asbestos has been studied by Pundsak [185, 186]. The void fractions observed experimentally were on the order of 6%. The water vapor adsorption-desorption data demonstrated that, in all of the samples, more than 80% of the void volume is in pores with radii less than 30 Å. Similar experiments have been carried out by Naumann and Dresher [187] with the use of nitrogen adsorption-desorption. A direct observation of porous structure in asbestos with the use of high-resolution electron microscopy was done by Yada [180, 188]. It was found that the outer diameter of the microfibers ranges from about 150 to 1000 Å, and the average value is about 500 Å. Most of the fibers were actually found to be hollow tubes with central holes 50–80 Å in diameter. The most frequent values of the outer and inner diameters of the samples showing a sharp distribution are 220–270 Å and 50–80 Å, respectively. None the tubes observed are simple cylindrical ones, but consist of spirally wound layers. Perfect concentric cylindrical layers were also found. The proportion of chrysotile fibers that are not hollow is very small. It was demonstrated that short uniform fibers of chrysotile asbestos exhibit strong interparticle forces parallel to the fiber axis, which result in aggregation of the fibers into distinct geometrical shapes [189].

Carbon Nanotubes The discovery of fullerenes provided exciting insights into carbon nanostructures. Carbon nanotubes are the most striking example. Carbon nanotubes were discovered by Iijima in 1991 during a study of the surfaces of graphite electrodes used in electric arc discharge [190]. Numerous novel and exceptional properties have been observed or predicted for these systems [191–200]. The basic structural unit of a carbon nanotube is a graphitic sheet rolled into a cylinder, with the tube tips closed by hemispherical or polyhedral graphitic domes. The aspect ratios of nanotubes vary with diameter, but the average length may be several micrometers. Experimentally, carbon nanotubes can be classified into two types: single-walled carbon

nanotubes (SWCNTs) and multiwalled carbon nanotubes (MWCNTs). SWNTs consist of single graphene cylindrical walls with diameters ranging between 1 and 2 nm. MWNTs have thicker walls consisting of several coaxial graphene cylinders separated by a spacing of 0.34 nm, which is close to the interlayer distance in graphite. The outer diameters of MWCNTs range between 2 and 25 nm, and the inner hollows, between 1 and 8 nm. Individual SWCNTs have a uniform diameter, although, when formed, they also show a strong tendency to pack together in large bundles. The very long cavities of carbon nanotubes have a special potential due to their high aspect ratio, and they can be used as templates for the fabrication of elongated nanostructures [201–205]. It may be noticed that there are some artificial noncarbon nanotubes based on MoS_2, WS_2, $NiCl_2$, etc. (see [206–210]). Some properties and the main methods for the synthesis of noncarbon nanotubes, as the results of experimental and theoretical studies, are discussed in [211].

3. METHODS FOR THE PREPARATION OF NANOSTRUCTURES WITHIN NANOPOROUS MATRICES

3.1. Filling with Wetting Liquids

Filling of nanovoids in porous matrices with various substances can be done from the molten state. In this situation, the possibility of filling the nanovoids strongly depends on such a characteristic of a system formed by the molten substance and the nanoporous matrix as wettability (see [212, 213]). If a liquid substance wets the inner surface of nanovoids it penetrates into these nanovoids and can organize inside the nanoporous material nanostructures whose shapes and sizes reflect the geometry of the porous space. So, to prepare nanostructures of this kind, we need in the simplest case only to melt some substance wetting the nanoporous matrix and to put this nanoporous matrix into this melt. There are many examples of nanostructures prepared in this way. First of all it should noted that water very frequently wets porous materials (and may even condense inside the nanopores from air) and so can be studied as a nanostructure [214–222]. Simple liquids like N_2, O_2, H_2, Ar, Ne, and CO have been introduced into nanopores of mainly Vycor porous glass [223–229] as superfluid liquids ^{3}He and ^{4}He [230–242] (zeolites have long been used as molecular sieves for wetting liquids). Various liquid crystals have also been studied in the confined geometry of the nanoporous space (mainly in nanopores of Vycor porous glass) [243–253]. Nanostructures have been prepared from the wetting liquid state or wetting solutions from organic materials and polymers [254–262], ferroelectric materials $NaNO_2$ and KH_2PO_4 [263–265], and fullerenes [266–268]. Carbon nanotubes have also been filled under the action of capillary forces in molten media [201, 203, 269] (see also [204, 205]).

3.2. Filling with Nonwetting Liquids

To fill porous matrices with liquids that do not wet the inner surface of porous space, it is necessary to apply some external pressure, which depends on the surface tension of liquid material and on the effective diameter of nanopores and corresponds to the formula for the Laplace pressure, $P = 4\sigma/D$, where σ is the surface tension and D is the diameter. In preparation of nanostructures within porous matrices, we need to put a porous matrix into the melt of some substance and to compress this system up to the pressures corresponding to the Laplace pressure, when the liquid substance fills the pore space and forms a nanostructure reflecting the geometry of nanovoids (see Fig. 5). This method of nanostructure fabrication is similar to the process of mercury porosimetry and, in the case of mercury-based nanostructures, the fabrication process exactly coincides with such a porosimetry. The characteristic surface tensions and the needed pressures (P_c) are presented in Table 5 for some liquid substances introduced into nanoporous matrices.

When the external pressure is decreased after filling, the liquid nanostructures are unstable because of a nonwetting situation, and some breaks appear even within nanovoids with uniform shapes. The nanostructures are in this case similar to some sets of isolated nanopatricles. However, the probability of breaking is lower if the external pressure is decreased after the system is cooled to the temperatures below the melting point, and the nanostructures may be continuous in this case. Nanostructures may be fabricated by filling Vycor porous glass with various metals at low temperatures of melting. Nanostructures have been prepared from Hg and Ga metals [270–276], In metal [277–281], and BiPb alloy [282–284]. Regular systems of nanovoids in zeolites [130, 285–288] and opals [153, 289–292] have also been filled with metals and other materials. Nanotubes of chrysotile asbestos have been filled under high pressure with Hg [293–297], In [298], Sn [299], Bi [300], Te [301], and InSb [300–303].

For this type of nanostructure preparation, there is an interesting question: What is the minimum diameter of nanopores for which this method can be used? This question is associated with such fundamental problems of capillary phenomena as determination of the local structure and properties of interface layers (so-called Tolman length

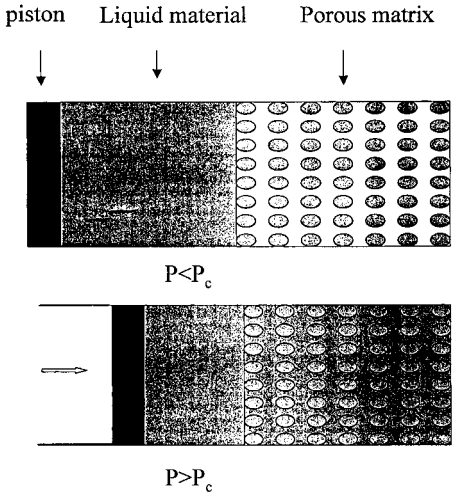


Figure 5. Process of porous matrix filling with molten material under high pressure.

Table 5. Surface tensions and typical needed pressures for filling of porous matrices with nonwetting liquids.

Liquid material (melt)	Surface tension (Erg/cm^2)	P_c (kbar, for $D = 100$ Å)	P_c (kbar, for $D = 20$ Å)
Bi	372	1.49	7.45
Hg	465	1.86	9.3
Pb	440	1.76	8.8
Se	105	0.42	2.1
Te	179	0.72	3.6
Sn	530	2.12	10.6
In	556	2.22	11.1
Ga	706	2.82	14.1

[213]), static and dynamic properties of the interphase interaction [212], capillary phenomena in extremely thin channels, or transition from capillary to diffusion phenomena. The experimental results from filling, with liquid metals, of channels of atomic-scale diameters in zeolites under high pressure [130, 286, 304–306] demonstrated that the interface transition layer has a thickness of about 1–1.5 Å, and the formula for the Laplace pressure can be used even for diameters of about 10 Å, in which this pressure approaches a value of $(20–30) \times 10^3$ atm. Capillary effects in this situation are a direct manifestation of the transformation of the mechanical energy (mechanical work done by a piston pressing some nonwetting liquid into a porous matrix) to the surface energy of a material highly dispersed in the pore space. Zeolites that contain liquid metals in their channels are heterogeneous systems with very high surface energy (one can estimate the energy capacity to be up to 200 kJ/liter) and can be used for simple direct storage of mechanical energy [305–309]. This high energy capacity (roughly 10^3 times greater than that of a spring) corresponds to that of an electrochemical storage unit, but, unlike the latter, it does not require an electric motor to produce mechanical work. Hydrocapillary storage units are sources of high-pressure liquid and can be used in transport, for example, with highly efficient hydraulic motors with high specific power (about 0.1 liter/kW) [307].

3.3. Chemical Methods for Filling of Nanoporous Materials

When some substance cannot be introduced into nanopores from a melt (such a situation is typical of materials with a high melting point), there are a lot of opportunities to prepare nanostructures through some chemical reactions inside the nanoporous space filled with various solutions (or vapors) of the components needed for such a reaction. Two main features characterize this method. First, it is not universal in comparison with the methods based on capillary phenomena and, therefore, requires in any particular case some special chemical approach that may be different from those employed. In the conventional bulk situation (a chemical reaction in the confined geometry of a nanoporous space may be not the same as in bulk). Second, such a method does not permit filling of the entire porous volume, because it is necessary to remove products of chemical reactions, and, as a result, the filling factor may be lower. A variety of chemical methods are based on electrolysis in an electrolyte introduced into the nanoporous system.

Metal-organic chemical vapor deposition (MOCVD) has been used to prepare GaAs and InAs nanoparticles from organogallium and organoindium compounds in Vycor porous glass [310, 311], InP nanoparticles in Vycor glass [312], and InP nanowires from trimethyl indium within asbestos nanotubes [313]. Other semiconducting compounds have been prepared in porous glass [314, 315]; CdS, CdSe, PbS, and PbSe compounds have been synthesized with the photochemical method [316]. The antiferromagnetic material MnO [317], like Fe compounds [310], has been introduced chemically into porous glass. Three-dimensional regular systems of Si and Si-Pt [318–322], GaN, InN, InGaN [323–326], and VO_2 [327] nanostructures have been fabricated in the void sublattice of a synthetic opal. The idea of using zeolites and mesoporous materials for the growth of quantum dots and wires has been proposed [328], and different nanostructures were synthesized within nanovoids of such porous materials as metals [329, 330], nanostructured metal sulfides [331], Ge [332], noble metal [333] nanowires, CdS nanorods [334], oxides [335], and organic materials and polymers [336–339]. Chemical reactions inside carbon nanotubes are described in [204, 205]; in particular, silver nitrate ($AgNO_3$) has been introduced (melting point 212 °C) into single-wall carbon nanotubes and decomposed into pure silver by thermal treatment at 400 °C, and the transformation of $AuCl_3$ in multiwalled carbon nanotubes into its base metal was described as an example of observation of reduction as a transformation of NiO to nickel metal. Similarly, encapsulated crystals of pure Pd [340, 341] have been prepared inside carbon nanotubes. Apart from hydrogen reductions, other kinds of reactions have been carried out inside carbon nanotubes. An example is the formation of cadmium sulfide crystals with hydrogen sulfide at 400 °C [342]. Electrodeposition has mainly been used for filling of some nanoporous membranes. Bi and Pt nanowires have been prepared in nanoporous anodic aluminum oxide [343–345] as Fe and Co nanowires [346] within nanoporous membranes and Ni nanowires in a porous Si matrix [347].

4. PROPERTIES OF NANOSTRUCTURES BASED ON NANOPOROUS MATRICES

4.1. Optical Properties

Artificially structured nanomaterials within porous matrices have attracted interest as objects of an optical study, because of the opportunity to observe quantum confinement, such as molecular-like discrete spectra, which exhibit strong size dependence, and some other effects, not associated with quantum confinement, which may be useful for different applications.

4.1.1. Quantum-Confinement Effects

Theoretical descriptions of the electronic structure of such nanostructures have advanced from considering a simple particle in box models to much more sophisticated approaches, including finite well effects, Coulomb interactions, nonparabolicity of the conduction band, confinement-induced mixing of valence subbands, etc. (see [348]). The simplest approach is based on the model of a spherical

quantum well with an infinite potential barrier, assuming a parabolic dispersion of electron and hole bands. The carrier energies are quantized and can be determined for semiconductors from the expression

$$E_{l,n}^{\mathrm{e,h}} = \left(\frac{h}{2\pi}\right)^2 \phi_{l,n}^2/(2m_{\mathrm{e,h}}a^2) \tag{11}$$

where $E_{l,n}^{\mathrm{e}}(E_{l,n}^{\mathrm{h}})$ is the electron (hole) energy reckoned from the bottom of the conduction (valence) band, $\phi_{l,n}$ is the nth root of the first-order spherical Bessel function, $m_{\mathrm{e}}(m_{\mathrm{h}})$ is the electron (hole) effective mass, and a is the nanoparticle radius. The absorption spectrum consists of narrow peaks with spectral positions

$$\left(\frac{h}{2\pi}\right)\omega_{l,n} = E_{\mathrm{g}} + \left(\frac{h}{2\pi}\right)^2 \phi_{l,n}^2/(2m_{\mathrm{r}}a^2) \tag{12}$$

where $m_{\mathrm{r}} = m_{\mathrm{e}}m_{\mathrm{h}}/(m_{\mathrm{e}} + m_{\mathrm{h}})$ and E_{g} is the energy gap for the bulk material. The absorption edge, which can be designated the energy gap for a nanoparticle, corresponds to the lowest transition and is shifted in comparison with the bulk situation. Because of quantum confinement, the gap increases as a^{-2} with decreasing nanoparticle size.

The first three roots $\phi_{l,n}$ are π $(l = 0,\ n = 1)$, 4.49 $(l = 1,\ n = 1)$, 5.76 $(l = 2,\ n = 1)$. So the absorption edge is given

$$by\left(\frac{h}{2\pi}\right)\omega_{01} = E_{\mathrm{g}} + \left(\frac{h}{2\pi}\right)^2 \pi^2/2m_{\mathrm{r}}a^2 \tag{13}$$

.

Because in semiconductors the hole mass is typically much greater than that of the electron, the electron levels are much more sparse than the hole levels. Estimation of the absorption edge shift gives a value of about 25 meV for $a = 5$ nm and $m_{\mathrm{e}} = m_{\mathrm{h}} = m_0$ (m_0 is the free electron mass). So the absorption edge shift due to quantum confinement can be observed in nanostructures within nanopores only a few nanometers in diameter or in semiconductors with sufficiently small effective mass of carriers. Luong and Borelli [349] have shown that III–V (GaAs, GaP) nanocrystallites in Vycor porous glass, introduced by metal-organic chemical vapor deposition, exhibit size-dependent shifts in the absorption and emission spectra, indicative of the quantum confinement. Optical properties of GaAs confined in mesoporous matrix MCM-41 also were studied [350]. Exciton absorption and luminescence spectra of PbI_2 nanocrystals [351] incorporated into porous glass by a low-temperature method (from solution) also demonstrated a short-wavelength shift due to quantum confinement. Indium phosphide has been deposited in 4-nm and 15-nm Vycor porous glass by the reaction of trimethylindium with excess PH_3 [312]. The absorption spectra of InP in a 4-nm porous glass have an onset at approximately 750 nm and reach a maximum at 500 nm. This sample is clearly blue-shifted with respect to the 15-nm sample, which displays a more broadened spectrum with an onset well before 1000 nm and a maximum at 700 nm. For comparison, bulk InP would show a sharp transition near 918 nm. Such a behavior of the absorption edge is consistent with quantum confinement. A broad absorption edge was attributed to a distribution of crystal sizes, since it has been observed that a 10% change in particle size results in broadened, featureless spectra.

A strong quantum confinement effect has been found for nanowires within chrysotile asbestos nanotubes, for GaAs [352–354], CdSe [354, 355], and InP [354]. The polarized absorption spectra of GaAs nanowires demonstrated bands at energies of 1.8 eV and 2.2 eV for both polarizations (the absorption is stronger for light polarized along the wire), the bandwidths being ~0.1 eV. Asbestos is transparent in the spectral region considered. Thus, all of the bands can be attributed to optical transitions in the GaAs nanowires. The energy of the first transition between the highest valence subband and the lowest conduction subband of a cylindrical quantum wire with infinite bandgap discontinuity can be estimated. If we take as a room-temperature energy gap $E_{\mathrm{g}} = 1.44$ eV for bulk GaAs, we obtain the quantum-wire energy gap and the first transition energy to be ~1.8 eV. This value correlates well with the spectral position of the 1.8-eV band. The band at 2.2 eV can be assigned to a transition between the valence band subbands split off by the spin-orbit coupling and the first conduction subband. The polarized absorption spectra of CdSe nanowires also display stronger absorption for light polarized parallel to the wires and weaker absorption for light polarized perpendicular to the wires. The broad band peaked at about 1.87 eV was observed and can be attributed to intersubband transitions in the cylindrical CdSe quantum wires. This band displays a blue shift of about 0.14 eV relative to the bulk CdSe absorption edge (the energy gap of the bulk CdSe is 1.73 eV at room temperature), which corresponds to the quantum-confinement energy shift of the first intersubband transition (for $m_{\mathrm{e}} = 0.12m_0$, $a = 3.5$ nm).

4.1.2. Metallic Composite Medium

The optical properties of metallic nanostructures within nanoporous matrices have been described mostly in the framework of the Maxwell-Garnett theory [356], which is a simple treatment of composite materials, addressing the shape but not the size of a particle. It provided the first theoretical basis to explain the resonant absorption (or dielectric anomaly), which is characteristic of systems of this kind. In the Maxwell-Garnett theory, the effective dielectric function $\varepsilon_{\mathrm{eff}}$ of a composite medium is related to the polarizability of a single metal particle via the Clausius-Mossotti equation [357], that is,

$$(\varepsilon_{\mathrm{eff}} - \varepsilon_{\mathrm{i}})/(\varepsilon_{\mathrm{eff}} + 2\varepsilon_{\mathrm{i}}) = (4/3)\pi N\alpha_0 \tag{14}$$

where N is the number density of metal particles, $\alpha_0 = R^3(\varepsilon_{\mathrm{m}} - \varepsilon_{\mathrm{i}})/(\varepsilon_{\mathrm{m}} + 2\varepsilon_{\mathrm{i}})$, ε_{i} is the dielectric function of the dielectric host, ε_{m} is the dielectric function of the metal, and R is the radius of the metal sphere. From this equation follows the relationship between the dielectric function of the metal-insulator composite and the dielectric functions of its constituents

$$(\varepsilon_{\mathrm{eff}} - \varepsilon_{\mathrm{i}})/(\varepsilon_{\mathrm{eff}} + 2\varepsilon_{\mathrm{i}}) = f(\varepsilon_{\mathrm{m}} - \varepsilon_{\mathrm{i}})/(\varepsilon_{\mathrm{m}} + 2\varepsilon_{\mathrm{i}}) \tag{15}$$

where $f = 4\pi^3 NR^3/3$ is the volume fraction of metal spheres.

In such a way the optical properties of a composite porous glass with Ag have been studied [358] by measuring the reflectance between 10,000 and 40,000 cm^{-1}. It was shown for samples with different sizes (mean radius from about 2.5 nm up to about 45 nm) that the red shift, broadening, and reduction of the peak in reflectance can be successfully compared with theoretical predictions. In a similar way nanoscopic gold cylinders of controlled radius (from 30 nm to 60 nm) and aspect ratio were studied; these were prepared via electrodeposition of the metal within the pores of anodically grown porous aluminum oxide membranes [359]. Maxwell-Garnett's original work addressed spherical inclusions, and it is for this specific geometry that the factor of 2 appears in the denominators. This factor is also called the screening parameter and, in general, is denoted by k (see, [360]). Hence, while $k = 2$ for spheres, long needlelike particles with their axes of revolution aligned with the direction of light incidence have a k value approaching unity. On the other hand, k tends to infinity for thin flat disks with an axis of revolution perpendicular to the direction of light incidence. The Maxwell-Garnett-calculated spectra are in qualitative agreement with experimental data for Ag in porous glass, which show dependence on both size and shape. Polarized optical spectra of long nanowires made of Hg, Bi, and InSb have been studied in the spectral range 0.5–3 eV [300]. These wires were prepared within channels of chrysotile asbestos by filling the channels with molten material under a pressure of up to 10 kbar. A high absorption anisotropy has been observed. This was explained in terms of the Maxwell-Garnett theory with depolarization factor k dependent on the nanowire orientation with respect to the electric field vector of incident light.

4.1.3. Interaction Between Filler and the Porous Matrix

In the case of nanostructures within porous materials, based on wetting liquids, the optical properties have attracted much attention because of the possibility of making some important conclusions about the phase state of encapsulated molecules and the possible mechanism of interaction between the filler and the active centers of the matrix, situated on the inner surface of pores, such as the molecular mobility of liquids within the nanopores. In such a way, the reorientational and vibrational dynamics of ethylene glycol and its homologous systems, possessing different numbers of OH groups per molecule, have been studied [257, 258] (see also [361]). The above systems have been studied in bulk and in the confined state within a porous glass matrix with pores 2.5 nm in diameter by means of light scattering (Raman and Rayleigh). By a comparison of the results obtained, evident influence of the chemical and physical traps on the molecular mobility was shown. The infrared absorption and Raman measurements, performed in the O-H stretching region, allowed identification of the intramolecular, H-bond subband and related various subbands to different intermolecular environments originating from the existence of the H-bond potential [362]. The optical properties of the 4-*n*-alkyl-4′-cyanobiphenils (*n*CBs, where n is the number of carbon atoms in the alkyl radical $n = 5, 8$) confined to porous glasses with four different mean pore sizes (2, 4, 54, and 90 nm), have been studied [255]. It has been shown that the interaction of *n*C8 molecules with the pore surface is accompanied by the formation of hydrogen bonds between the SiOH group on the glass pore surface and the cyano group of the *n*CB monomer. In contrast, in the case of benzophenone confined to porous glass, weak OH-Si hydrogen bonds are formed. Phosphorescence spectra of benzophenone encapsulated in the inner space of mesoporous aluminosilicate molecular sieves of the MCM-41 type (the channel diameter is about 4 nm) have been analyzed at 4.2–200 K [254]. A comparison of these spectra with the spectra of a bare MCM-41 matrix, pure benzophenone, and benzophenone incorporated into the porous glass led to conclusions about the phase state of encapsulated molecules and the possible mechanism of interaction between the filler and active centers (of the Si-OH type) on the inner surface of pores. The main influence exerted by the strong dipolar interactions on the dynamics of hydrogen-bonded liquids (propylene glycol) diffusing within 2.5-nm and 7.5-nm pores of a sol–gel porous glass has been demonstrated [363] by depolarized light-scattering spectroscopy. The experimental spectra measured as a function of temperature revealed unambiguously the slowing down of the collective reorientational processes triggered by the confinement effects. In addition to intermolecular interactions, an interaction mechanism of induced absorption has been identified in porous Vycor glass filled with liquid H_2 by measuring the infrared absorption [364]. Infrared spectra have been studied for pyridine adsorbed on porous glass [365], and it was established that physical adsorption occurs by hydrogen bonding of pyridine to SiOH and B-OH groups via a nitrogen atom. Absorption in the far infrared region has been used to investigate surface vibrations in NaCl crystals dispersed in porous glass [366]. The experimental results were in good agreement with the theory predicting a gap between the limiting frequencies of the longitudinal and transverse optical surface vibrations in very small ionic crystals. Optical methods have also been used to study the interaction between liquid crystals and the inner surface of porous matrices [250, 252, 367, 368]. Dynamic and static light-scattering measurements in a nematic liquid crystal, pentylcyanobiphenil (5CB), confined in silica porous glasses with an average pore size of 100 nm (volume fraction of pores 40%) and 10 nm (volume fraction of pores 27%), demonstrated significant changes in the physical properties of confined liquid crystals [252]. The nematic-isotropic phase transition temperature is depressed by 0.6 K in 100-nm pores, compared with the bulk value, and this phase transition is not detected at all in 10-nm pores. It was found that even at about 20° below the bulk melting point the relaxation processes in a confined liquid crystal are not frozen. The infrared absorption spectra of 5CB dispersed in porous glass and molecular sieves of the types Si, Al-MCM-41 and Cu-Si, Al-MCM-41 (pore diameter of about 4 nm) have been investigated in the region of fundamental and overtone vibrations [367]. The mechanism of interaction between a part of 5CB molecules and the active centers of the porous interface was estimated. In this case, a strong enough interaction occurs between the cyano groups of 5CB and the hydroxy groups of the channel surface. It was also shown that another part of the 5CB molecule exists in the liquid-crystal state, and its

amount depends on the size and nature of the pores. Photon correlation spectroscopy was used to investigate nematic liquid crystals dispersed in porous matrices with randomly oriented interconnected pores (porous glasses) and parallel cylindrical pores (Anopore membranes) [250]. Investigation of liquid crystals in cylindrical pores, together with studies in random porous matrices, made it possible to separate the effect of random structure and domain formation from the contributions due to the existence of a liquid crystal–solid pore wall interface and the effect of the finite pore size in relaxation of the order parameter or director fluctuations.

4.1.4. Luminescence, Phosphorescence, Spectral Kinetics

Interface phenomena in nanostructures formed within porous matrices have also been studied with the use of luminescence, phosphorescence, spectral-kinetic techniques, etc. (see [266, 267, 313, 369–376]). An enhanced and broadband photoluminescence behavior of fullerene (C_{60})-doped porous glasses has been observed [371]. By means of photoluminescence and photoluminescence-excitation spectroscopy of C_{60} dopant in surface-unpassivated and -passivated porous glasses, it was shown that the interaction of C_{60} with the silica surface is quite strong (see also [267]) and that the modified C_{60} molecules contribute only in the orange-red region (1.9 eV) of the observed photoluminescence spectrum. The luminescence spectra of a number of polypyridine ruthenium (II) complexes adsorbed on the surface of a porous glass have also been studied [372]. Experimental data on the decay of the excited state of adsorbates were adequately described within the framework of a two-exponential model. Special features of the spectral luminescence parameters measured were accounted for by inhomogeneous broadening of spectra of luminophores arranged under heterogeneous conditions on the glass surface. The revealed features of luminophore quenching in heterogeneous conditions [373] were explained within the framework of a model analyzing two subsystems of adsorbed luminophores. Different mechanisms of quenching were considered, and a conclusion about the dynamic character of luminophore quenching in a porous glass was made. A comparative study of Raman, optical absorption, and photoluminescence spectra in 3D regular ensembles of InP quantum wires [313, 369, 377] revealed the dependence of the optical properties of these quantum wires on interfacial effects, namely, atomic interactions in the wires, and wire-matrix and wire-wire interactions. InP quantum wires have been produced in channels of chrysotile asbestos (wire diameter $D = 8$ nm, distance between wires $A = 40$ nm), MCM-41 mesoporous framework silicate ($D \approx 3$ nm, $A \approx 3.4$ nm), and $AlPO_4$-5 zeolite ($D \approx$ nm, $A \approx 1$ nm) by a two-stage gas-phase substitution reaction of the metal-organic compound trimethylindium [378]. InP wires in chrysotile asbestos are virtually isolated from one another; in other words, there is no wire-wire interaction. The spacial separation between the channels in MCM-41 is small enough (channel wall thickness of 0.8 nm) to allow effective coupling between nanowires. Taking into account that only one InP molecule can fit to the diameter for $AlPO_4$-5 zeolite, it appears natural to assume that the interaction between the $AlPO_4$-5 matrix and InP is the strongest among all of the hosts considered. A comparison of optical data for these three matrices demonstrated that the wire-matrix interaction distorts the InP lattice, broadens the electronic density-of-states spectrum of the wire in the vicinity of the energy gap, and redistributes the relaxation of photoinduced excitations among states belonging to the wire itself and to defects in the matrix bound to the wire.

The migration of triplet excitations of complex molecules in a confined geometry was studied in [374–376]. The results of investigation of the decay kinetics of phosphorescence and delayed fluorescence of disordered chrysene and chrysene in porous matrices were presented. The study of triplet-triplet annihilation kinetics allows an understanding of transport in disordered films, polymers, and membranes and of paradigms of heterogeneous chemical kinetics. It is also a tool for studying the topology, morphology, and structure of molecular aggregates, strands, pores, and domain boundaries. Porous glasses represented as chaotically arranged spheres and natural mineral chrysotile asbestos represented as hollow cylinders packed parallel to one another were used as matrices. An investigation of delayed luminescence of disordered chrysene and chrysene in porous matrices has shown that the triplet excitation transport in these systems at $T = 77$ K is dispersive in nature. The annihilation kinetics of triplet excitations is described by a rate coefficient that depends on time according to a power law. The exponent is determined by the microscopic structure of a sample and its geometry. In the case of chrysene in porous matrices, the annihilation kinetics of triplet excitations is determined by the matrix geometry and varies slightly with increasing temperature. An analysis of the decay kinetics of delayed luminescence from chrysene in porous glasses in comparison with chrysotile asbestos has led to a conclusion that the topology of the porous network of porous glasses is similar to that of the 3D percolation cluster.

4.1.5. Structure of Clusters and Nanochains in Zeolites

Zeolites, which possess a regular system of cavities with a diameter of about 1 nm, are very appealing, in that they permit the preparation of arrays of unisized nanoclusters in the cavities. There are a lot of questions about the kind of structures that may be stabilized inside the zeolite pores. Some of these have been solved by applying optical spectroscopy methods to clusters of chalcogens (sulfur, selenium, tellurium) introduced into the pores of zeolites NaA and NaX (see [319, 379–387]). Chalcogens are good materials for the preparation of this kind of cluster, because they can easily be introduced from the melt or adsorbed into zeolite nanovoids, and their optical properties can be studied in the visible and near-UV spectral ranges, in which zeolites are transparent. For sulfur clusters it has been shown by Raman scattering and optical absorption spectra [382] that the main molecular unit of sulfur in the NaA zeolite is the most stable sulfur molecule, namely, the S_8 crown-like ring. NaA with selenium has been examined by Raman [380, 381, 383, 384, 386, 387] and optical absorption [379, 381, 385–387] spectroscopies. The whole body of data obtained

shows that some kind of Se ring molecules, rather than chain molecules, are stabilized. The NaA zeolite with Te has also been studied by Raman [380, 386] and optical absorption [386] spectroscopies. Similar to the case of Se clusters, the experimental data obtained indicate that, probably, Te ring molecules are stabilized. A more detailed study of Raman spectra [388] yielded evidence in favor of the stabilization of S_8, Se_8, Se_{12}, and Te_8 ring molecules in the NaA zeolite. S_8 and Se_8 molecules are well known to exist in the condensed state, but there has been no information about the stabilization of Se_{12} and Te_8 rings under conditions different from those in the zeolite NaA cavities. A possible reason for this stability is the good compatibility of the size and symmetry of Se_{12} and Te_8 with the size and symmetry of cavities in the NaA zeolite. Strong and broad low-frequency bands in the Raman spectra of a NaA zeolite with sulfur and selenium were attributed to librations of ring molecules in the cavities. To obtain information about the symmetry of the Raman-active vibrations and the orientation of clusters in zeolite cavities, Raman microprobe polarization measurements have been performed (usually microcrystalline powder samples were studied) [389]. The dependences obtained correspond to S_8 rings with the D_{4v} point group symmetry, with their fourfold axes aligned with the fourfold axes of zeolite NaA, and Se_{12} rings with the D_{3d} point group symmetry, with their threefold axes oriented along the threefold axes of the zeolite. The possibility of determining the ratios between the Raman tensor components of vibrations of zeolite-confined clusters was demonstrated.

There are zeolites, such as mordenite, cancrinite, and $AlPO_4$-5, which are very attractive for the preparation of 1D systems. The chalcogen guest materials (sulfur, selenium, tellurium) have been introduced into the cavities of these zeolites, and the structure of chain-like nanostructures within the pores has been studied by optical methods (see [381, 384–386, 390] for mordenite matrix, [391–393] for cancrinite matrix, and [394–396] for $AlPO_4$-5 matrix). Polarized Raman and optical absorption spectra of natural mordenite single crystals containing sulfur, selenium, or tellurium within 1D channels (channel diameter 0.7 nm) have been studied. The bands in the Raman and absorption spectra, which are polarized along the channel direction, are attributed to helical chains. The helical chains and ring molecules (S_6, S_8, Se_6, Te_6) coexist in the mordenite channels. The chains contribute to absorption spectra only for the E$\|$c polarization. Correspondingly, the resonance enhancement of Raman spectra of the chains is working for cc-polarization, and so Raman bands of the chains are active in this polarization. Ring molecules absorb light for all possible polarizations, and so the molecule vibrations are active for the aa-, bb-, and cc-polarizations. The ratio between the concentrations of chains and rings can be related to some kind of regular arrangement between chains and rings inside a channel.

Polarized Raman, optical absorption, and luminescence spectra of cancrinite crystals possessing parallel nanochannels with Se inside have been investigated. It was shown that Se is stabilized in the form of Se_2^{2-} and Se_2^- dimers situated at the center of the channel and oriented along the channel. When present in high concentration, the Se_2^{2-} dimers show a tendency to organize linear chains. At low temperatures, quite strong interdimer bonding for both Se_2^{2-} and Se_2^- was observed. The additional Raman bands were attributed to vibrations of linear Se_2^{2-} chains distorted by the incommensurate potential of cancrinite. Strong near-infrared polarized luminescence was observed. Photoionization of dimers was shown to be an important step in the mechanism of the luminescence.

From Raman spectra of selenium incorporated into the channels of $AlPO_4$-5 zeolite single crystals ($AlPO_4$-5 is a molecular sieve with nearly cylindrical channels with diameter of 0.73 nm) it has been concluded that single helical chains and Se_8 molecules are formed inside the channels. The temperature dependence of the Raman spectrum of single helical Se chains was studied. Dramatic changes in the spectra owing to structural transformations in this 1D system were observed. The chains are ordered at 77 K. One-dimensional Se shows transformations of the chains from the ordered state at low temperatures to weakly and strongly disordered states at higher temperatures. A phase transition from a weakly to a strongly disordered state is accompanied by structural relaxation of the chain. The structural transformations observed in single Se chains can be regarded as an example of how a nearly 1D system confined in a nanochannel behaves.

4.1.6. Nonlinear Optical Properties

The potential of quantum-confined semiconductor nanoparticles for nonlinear optical and electro-optical applications has been anticipated [397–399]. Experimentally, the nonlinear optical properties of nanoparticles in porous glass have been studied for GaAs [400, 401], InP [402], and Cu_2O [315]. Time-resolved pump/probe Z-scan experiments (for the Z-scan technique, see [403]) using 1064-nm 100-ps pulses have been performed to determine the temporal characteristics of the optical nonlinearities of GaAs nanocrystals. The buildup and decay of the nonlinearities were instantaneous, compared with the duration of the excitation pulses. Single-beam Z-scan experiments were also used to measure the two-photon absorption coefficient, β, and bound electronic nonlinear refractive index, γ, of the quantum-confined crystals. The ratio γ/β is greater by approximately a factor of 10, compared with bulk GaAs. In the case of InP nanocrystals, the magnitudes of nonlinearities in a 15-nm sample were found to be similar to those in bulk InP, whereas for a 4-nm sample the ratio γ/β was a factor of 14 greater, compared with that in bulk InP.

Carrier confinement to a single dimension must also lead to enhanced optical nonlinearities [397], and, therefore, the nonlinear optical properties of quantum wires have attracted much attention. Nanowires in channels of chrysotile asbestos (GaAs, CdSe, InP) have been intensively studied in this regard [354, 404–411]. The experiments used the pump and probe method to measure the differential transmission $\mathrm{DT}(\omega) = [T(\omega) - T_0(\omega)]/T_0(\omega)$, where $T(\omega)$ and $T_0(\omega)$ are the transmission spectra of the excited and unexcited samples. Nonlinear optical absorption at discrete frequencies (bleaching bands) has been observed for GaAs quantum wires with an average diameter of about 6 nm. The induced decrease in absorption was accounted for by filling of the quantum-well energy levels

of quantum wires with nonequilibrium carriers (saturation effect). Strong ($\chi^{(3)} \approx 10^{-8}$ e.s.u.) and fast (relaxation time $\leq$50 ps) third-order optical nonlinearity has been revealed. The nanowires within chrysotile asbestos demonstrate pronounced excitonic behavior with enhancement of excitons, with the binding energies ranging from 120 to 260 meV.

4.1.7. Photonic Crystals

The concept of a photonic crystal whose behavior with respect to photon waves is similar to that of a dielectric crystal with respect to electron waves has been advanced by Yablonovich [412] and John [413]. In periodic dielectric structures with a period close to the wavelength of electromagnetic waves, these waves undergo Bragg diffraction, and a stop band is formed for modes propagating in a given direction. If the stop bands for all propagation directions overlap in some frequency range, a photonic gap can be created, in which the density of photonic states is zero. It has been shown that synthetic opals [414–417], like opal-like films formed from polymethyl methacrylate beads [418], possess the properties of photonic crystals throughout the visible spectrum. Moreover, because of their regular porous structure, opals can be used as matrices for the fabrication of regular arrays of nanosized systems in the voids between the spheres constituting opal, which may also possess photonic stop bands. The voids between the silica spheres may comprise up to 26% of the total volume of the opal material. This makes it possible to change the optical contrast coefficient ($\eta = (\varepsilon\nu/\varepsilon_s)^{1/2}$, where $(\varepsilon_s)^{1/2}$ and $(\varepsilon\nu)^{1/2}$ are, respectively, the bulk refractive indices within the SiO_2 spheres and outside them [419]) by filling the voids with different substances. According to the theoretical estimations [419], a complete photonic bandgap may exist at $\eta \geq 2.8$. In pure opals, where η is considerably lower, the formation of the complete photonic bandgap is, apparently, impossible. Therefore, materials with a high dielectric constant should be used to fill the voids. In a similar way nanosystems of CdS particles [420] have been prepared: GaN, InN, and InGaN particles [323–326]; Si particles; VO_2 particles [327]; and photonic crystals made of nematic liquid crystals [421]. Moreover, after filling of the original opal matrix, SiO_2 can be removed chemically, and opal replicas can be prepared from different materials [422–424]. This operation yields a 3D semiconducting lattice, which occupies up to 26% of the total volume of the material and is surrounded by a "matrix" of air spheres that occupy the remaining 74% of the volume. The potential applications require control of the photonic gap structure in real time under external effects. It has been proposed [421, 425] that the opal pores be filled with liquid crystals, which allow tuning of characteristics of the photonic bandgap structure through an electro-optical effect in the liquid crystal, occurring when an electric field is applied. It has been demonstrated [327] that the photonic bandgap structure can be controlled via the metal-semiconductor phase transition in VO_2 embedded in opal. A controllable and reproducible variation of the photonic bandgap properties of the opal-VO_2 composite during heating and cooling in the temperature range 60–90 °C was achieved. This is due to the change in the dielectric constant of VO_2 upon the phase transition. It was suggested that a phase transition in this material and, hence, the photonic bandgap tuning in the opal-VO_2 composite can also be observed under ultrafast laser pulses.

4.1.8. Laser Media Based on Nanoporous Materials

New optical media have been developed on the basis of nanoporous materials, mainly porous glasses, for both optoelectronic applications [426] and laser elements [427]. New types of lasers have been fabricated by the incorporation of active dyes into porous glasses [428–434]. Molecular sieves can host some laser active dyes, too. Pyridine-2 molecules have been introduced into the channel pores of nanoporous $AlPO_4$-5 zeolite [435]. It was established that many effective dye molecules, such as rhodamines, do not fit into the pores, but the amount of encapsulated dyes can be raised by modifying the structure of the dyes so that they could match the host templates. The resulting microlasers have properties that depend on the size and shape of microcavities. For dyes that fit into the pores, partial regeneration of the photoinduced damage has been observed.

4.1.9. Materials for Optical Data Recording

Filling of porous media with photosensitive materials has been used to prepare new media for optical data recording (silver-containing porous holograms). With these materials, one can make a volume holographic record with high diffraction efficiency [436–442]. Deep 3D holograms with large physical thickness, allowing postexposure amplification and a posteriori alteration of the grating parameters, were obtained. The sol–gel technology for making glass has made it possible to synthesize a new class of porous materials, monolithic xerogels. Xerogels have a more uniform porosity, which is an important advantage in data recording. It has been shown that holograms with a diffraction efficiency of nearly 100% can be obtained upon an exposure of about 0.1 J/cm.

4.2. Electronic Transport

4.2.1. Superconductivity

Superconducting Properties of Metals in the Confined Geometry of Porous Glass The first studies of superconductivity in confined geometries of porous glass were concerned with the development of new artificial hard superconductors [443, 444]. New superconducting nanostructures were fabricated by pressing mercury into porous Vycor glass placed in mercury in a small balloon and compressed hydrostatically for 20 min at 3300 atmospheres, the pressure required to force the mercury into the Vycor glass. By measuring the weight gain, it was found that 10% of the total volume of the sample, whose interconnected pores were approximately 3 nm in radius, was filled with mercury. The superconductivity was measured with the magnetic induction technique. The measurements of magnetization and critical current demonstrated a large enhancement of the critical magnetic field and critical current in comparison with their bulk values. The critical current density was $\sim 10^4$ amps/cm^2, and the critical magnetic field, $\sim$40 kOe at 2.1 K (the critical magnetic field for bulk mercury at 2.1 K is 320 Oe).

This size effect may arise when the characteristic sample size is made comparable to, or less than, the bulk coherence length ξ_0 and the magnetic field penetration depth λ_0. For bulk mercury, $\xi_0 \approx 200$ nm and $\lambda_0 \approx 43$ nm, and, therefore, the observed enhancement is not surprising. Furthermore, the superconductivity of some other metals (In, Sn, Pb, Tl, Ga) within nanopores of porous glass was studied [445–447]. Measurements of the critical magnetic field and transition temperature of superconducting indium in porous glass have been made with a low-frequency mutual-inductance technique [446]. Pore diameters from 6.5 to 25 nm were used. It was found in these measurements that the critical magnetic field H_c depends strongly on D, the pore diameter of the glass. Below $t = 0.5$ ($t \equiv T/T_c$, T_c critical temperature) H_c can be represented as approximately $H_c = (3415 \pm 40)(1 - t^2)/D^{(1.00 \pm 0.14)}$, where H_c is measured in kOe and D in Angstroms. This is in agreement with the predictions of H_{c2} made by de Gennes and Maki [448, 449] for type II superconductors in the dirty limit under the assumption that the electron mean free path is proportional to D. According to the de Gennes-Maki theories,

$$H_{c2}(0) = (3/2\pi^2)(\Phi_0/\xi_0 l) \tag{16}$$

where Φ_0 is the flux quantum, ξ_0 is the coherence length of the pure metal, and l is the electron mean free path. Near $t = 0$, de Gennes and Maki obtained

$$H_{c2} \cong 0.87\lambda_L(0)H_c(0)(1 - t^2)/l \tag{17}$$

where $\lambda_L(0)$ is the London penetration depth in pure metal at $t = 0$. The dependence of T_c on D is conveniently represented as $T_c - T_{cbulk} = 1 - 0.028/D$, where D is in Angstroms. It was supposed that the change in T_c may be due to strain, mean free path, or surface effects. Experimental results on the transition temperature of indium, thallium, and lead grains of various sizes in porous glass, together with preliminary measurements on other metals [447], have been interpreted in terms of changes in the phonon spectrum, which alter the electron-phonon coupling constant. McMillan [450] has shown that the transition temperature of simple metals can be raised by reducing the Debye temperature. It has been suggested [451, 452] that, in small crystallites, the phonon spectrum is appreciably modified by a greater proportion of low-frequency surface phonons, and this is the reason for the shifts in T_c in granular or disordered materials. For these purposes the phonon spectra of gallium and tin embedded in porous glass have been studied with the Mössbauer technique [453, 454]. In [281], superconducting metals within porous glass were described as granular superconductors, that is, as superconductors consisting of small grains that are separated spatially but connected by electron tunneling. Abeles et al. [455] have carried out a calculation for a 1D model of a granular superconductor, with the grain boundaries represented by δ-function barriers. They found that the material behaves like a homogeneous dirty superconductor, with a critical field given by de Gennes-Maki theory, but with an effective mean free path $l = D\tau/(1 - \tau)$, where D is the grain size and τ is the transmission coefficient for electrons incident on a δ-function barrier. Some experimental results and results derived from this theory for metals in porous glass are given in Table 6.

Table 6. Parameters of superconductivity for metals in porous glass.

	D (Å)	T_c (K)	H_{c2} (kOe)	ξ_0 (10^5 cm)	τ	l (Å)
Indium	31	4.24	69	3.5	0.042	1.31
	60	4.05	39	3.7	0.036	2.17
	80	3.96	29	3.75	0.036	2.90
Lead	32	7.05	96	0.85	0.120	3.84
	58	7.15	55	0.83	0.120	6.90
Tin	31	4.97	54	1.72	0.11	3.42
	59	4.25	39	2.00	0.104	4.40
Thallium	32	2.65	48	1.17	0.175	5.59

It has been shown that Pb-Bi alloys embedded in a porous glass [282, 283, 456] can be used in some applications because of their very large critical magnetic field (for Pb-40%Bi H_{c2} was about 120–125 kOe at 4.2 K). The transition temperature was about 7.8 K for porous glass with a 6-nm pore diameter. Measurements of the critical current density for Pb-40%Bi alloy in porous glass with an average pore diameter of 3.5 nm have been made in a transverse magnetic field and demonstrated a typical critical current density of about 10^5 A/cm^2 in a zero magnetic field and 10^4 A/cm^2 at 90 kOe. For Pb-Bi alloy in porous glass with a 2-nm pore diameter [456], the critical temperature was found to be 6.2 K, and $H_{c2}(0)$ was 230 kOe (as estimated by extrapolating from 210 kOe, the highest field used in measurements). Interestingly, the last value is greater that the so-called paramagnetic limit introduced by Clogston [457] and Chandrasekhar [458]. They pointed out that the electron Zeeman energy in the normal state can make a significant contribution to H_{c2} determined at high enough fields. At magnetic fields stronger than $H_p = [\Delta_0/(2\mu_B)^{1/2}$ (where Δ_0 is the energy gap of the superconductor and μ_B is the Bohr magneton) superconductivity must be impossible, since the free energy of electrons in the normal state is equal at H_p to the superconducting condensation energy. For PbBi alloy, H_p is about 120 kOe, and so the experimental value of H_{c2} for this material in porous glass exceeds the paramagnetic limit. The explanation was found in the framework of the theory [459], including the effects of spin-flip scattering induced by spin-orbit coupling.

Gallium has been introduced under a pressure of 14 kbar from a melt into porous media, such as cellulose, corundum powder, synthetic zeolites (NaA, NaX), and porous glass [271]. In the temperature range studied (4.2–16 K), two superconducting transitions were observed in the range from 6.1 to 6.4 K and between 6.8 and 7.2 K. The first transition was ascribed to the metastable phase β-Ga, and the second, to another metastable phase, named δ-Ga. These phases were previously observed in thin films and in gallium emulsions [460, 461] (T_c of the ordinary phase of gallium is 1.08 K). The major part of gallium in porous glass was not superconducting above 4.2 K. At 4.2 K, a broadened signal was observed, which could hardly be suppressed with a magnetic field. This signal was identified with the "tail" of some low-temperature ($T_c < 4.2$ K) transition of gallium into the pores. Further study of the superconductivity of gallium embedded in nanoporous materials (various porous glasses and opal) was performed [462–466] to reveal how the size and configuration of pores influence the superconducting

features, because it was found previously that many general superconducting properties of porous glasses filled with metals occur because of the interplay of strong and weak links between metallic nanoparticles in pores. The superconductivity was studied with a superconducting quantum interference magnetometer with a 7-T solenoid in the temperature range 1.7–20 K. Single and double superconducting phase transitions were observed for different samples. Magnetization hysteresis loops were also measured and found to be dependent on the size and geometry of the pores. The changes in magnetization below about 6.4 K were analyzed in terms of models for granular superconductors, while the alterations in magnetization near 7.1 K were regarded as resulting from a superconducting phase transition in a coexistent structural modification of confined gallium. X-ray diffraction measurements confirmed the presence of such an additional gallium modification. The results obtained suggest that studies of magnetization at low temperatures can be used to get information about the geometry of the pore network and distribution of Josephson links in porous composite materials.

A double resistive superconducting transition has also been found for indium within 5.6-nm-pore Vycor glass [467] (see also [468, 469]). The interconnected network of pores occupied approximately 30% of the total sample volume. From the broadening of the X-ray diffraction lines, a characteristic crystalline size of about 35 nm is inferred, which is significantly larger than the pore size of the Vycor. Furthermore, the X-ray diffraction measurements do not reveal a second phase. At the same time, the zero-field resistive transition shows a two-step temperature dependence with maximums in the derivative dR/dT at 3.99 and 4.03 K (T_c of bulk indium is 3.40 K). In a magnetic field, two fundamentally different transitions were observed by means of electric transport and magnetic susceptibility measurements. It was concluded that the presence of two distinct transitions is an intrinsic effect associated with the microstructure of indium in the porous Vycor glass. The superconducting transition temperature of an In-impregnated porous glass with 3.8-nm pore diameter has also been measured as a function of a hydrostatic pressure of up to 3.74 kbar [280]. All measurements were performed with a sample inside a Be-Cu pressure cell. The mean T_c increased above the value characteristic of bulk In by 17%. Its pressure derivative dT_c/dp was determined to be $(-2.8 \pm 0.6) \times 10^{-5}$ K/bar less than the effect of pressure in bulk indium. It was concluded that the increase in the critical temperature and the smaller effect of pressure on the critical temperature for In-impregnated porous glass, compared with bulk indium, are both due to the softening of the phonon spectrum, resulting from the increased surface-to-volume ratio of the pores.

Superconducting Properties of 3D Arrays of Weakly Coupled Nanoparticles within Opal Arrays of nanosize superconducting metal particles have been prepared by impregnation of the opal matrix with molten metal (In, Pb, Sn, Ga, and PbBi alloy) under high hydrostatic pressure [153, 289–292, 470]. The shape of the metal grains is an exact copy of the void configuration, since the metal occupies the entire free volume of the matrix, and so a 3D replica of the matrix can be formed. In accordance with the matrix structure [150], each section consisting of two adjoined grains contains a constricted region that can be considered a weak link (in the sense of superconducting properties), that is, an element of the S-c-S (S, superconductor; c, constriction) type, and so the proposed material contains a regular array of identical Josephson junctions. Owing to the matrix, these elements are arranged in a crystalline manner and form a macrosystem. The density of weak links in this material is about 10^{14} cm^{-3}. The interest in systems with a large number of Josephson junctions has been growing because of the practical demand. Microwave generators, receivers, and other devices based on a single Josephson junction are limited in their application because of their inability to emit powerful radiation in the generator mode and their liability to saturation in a weak external field in the detector mode. In-series connected Josephson junctions are very promising for application as power sources of microwave radiation for space communications, wide-dynamic-range detectors, parametric Josephson amplifiers, voltage standards, etc. Their main advantage is the coherent addition to the output signal from each element of an array. Microwave generators using coherent oscillations of an array of 40 or 100 Josephson junctions have been fabricated with good performance: emitted power of up to 10^{-6} W and frequencies of up to 300 GHz; moreover, the phase-locked emission linewidth was less than 5 kHz [471, 472]. In addition, an increase in the dynamic range of detectors in proportion to N^2 is observed for N in-series connected junctions [473]. The results obtained with opal-based Josephson junction arrays show that the materials prepared behave like continuum Josephson media exhibiting mutual phase locking. The following dependences were measured: resistance versus temperature, $R(T)$; critical current versus temperature, $I_c(T)$; and current–voltage characteristics. The measured detector response to an external radiofrequency signal were compared with $R(T)$ and dR/dT dependences and current–voltage characteristics. The microwave sources in use covered the range from 0.5 to 10 GHz. To separate the amplitude- and frequency-dependent response components from the total detector signal, amplitude-modulated radiation and frequency-modulated radiation were used. Magnetic properties of the system in a weak magnetic field $H < 150$ Oe, namely, the dependences $I_c(H)$ in dc and ac magnetic fields, shift of the superconducting transition temperature, and distortions of the voltage–current characteristics and responses, were also studied. The resistive transition demonstrated a shift toward higher temperature. For the In-based system, the critical temperature was shifted to $T\ (R = 0) = 3.446$ K, compared with that in bulk In ($T_c = 3.40$ K). Taking into account the experimental dependences of T_c on the diameter of superconducting nanoparticles, one can estimate the related size of In grains to be 50 nm. This value corresponds to the void size. The critical current density of each bridge was on the order of 10^4 A/cm^2, which is about 10 times less than the characteristic bulk value, and the sample-average critical current density was about 10^2 A/cm^2. An oscillating dependence of the critical current on the external magnetic field has been observed. Conventionally, such a dependence corresponds to trapping of flux quanta in the lattice cells of a phase

coherent array of Josephson junctions. For a Sn-based system the periods of the observed oscillations, $\Delta H_1 = 3$ Oe in a weaker magnetic field and $\Delta H_2 = 22$ Oe in a stronger field, were smaller than the critical field in bulk Sn ($H_c(0) =$ 306 Oe) at $T = 3.42$ K, fixed in the experiment. Therefore, this effect can be regarded as resulting from quantum interference in a regular lattice of magnetic vortices. These data were fitted using the relation $\Delta H \pi D^2/4 = \Phi_0$ (Φ_0 is the flux quantum). The first period corresponds to a lattice of loops with effective diameter $D = 2500$ nm. When a state with 4 flux quanta per loop is achieved, the system jumps to a state with loops 1000 nm in diameter. Both periods correspond well to multiples of the silica sphere diameter (250 nm). The typical voltage–current characteristics of the samples under investigation look like the curves observed for bridge-type Josephson junctions, which obey the resistively shunted junctions model. Typically, the hysteresis loop of the voltage–current characteristic takes place at $T < T_c$. However, other specific features of the voltage–current curve were found: switching between some fixed states in the resistive branch of the voltage–current characteristics; irreversibility of voltage–current characteristics (that is, any shift of the operating point along the resistive branch forward and back results in a hysteresis loop); an "anomalous" hysteresis loop at temperatures that are low but close to T_c, that is, the downward branch of voltage–current characteristics branch passes above the upward branch [153, 289, 290]. As for the last phenomenon, upon cooling the magnitude of this hysteresis first reaches a maximum and then gradually decreases and turns into the "normal" one. Moreover, the stronger the current through the sample ($I > I_c$), the greater the "anomalous" difference between the upward and downward branches of the voltage–current characteristics. The response to microwave radiation shows no Shapiro steps in the voltage–current characteristics of a sample exposed to external radiation; that is, there is no selective detection typical of separate junctions. However, the origin of detection of the radiofrequency signal is Josephson-related rather than classical. The frequency-modulated response curve differs significantly from the amplitude-modulated one both in magnitude and in shape; that is, the sample exhibits frequency-dependent properties. So, a Josephson-like behavior of such a system was demonstrated, but, at the same time, some disadvantages were found: the indefiniteness of its structure, related to the polycrystalline nature of artificial opal-like matrices, and the uncertainty of the type of interjunction interaction due to the close displacement of Josephson junctions.

Superconductivity in Metallic Nanowires The nanochannels of chrysotile asbestos matrix have been filled under high pressure with superconducting metals (Hg, In, Sn, Ga, Pb), and the superconducting transitions of a complete series of such samples have been studied by the contact method [293–299, 474, 475]. Relieving the external pressure after filling of asbestos channels with liquid metals led to a loss of sample conductivity because of some oozing of liquid metals from the pores due to the nonwetting instability. It was, therefore, essential to study the system directly under high pressure, using autonomous chambers where the pressure was kept. Each sample contained about 10^6 parallel nanowires with the same diameter and length of about 0.5 cm, separated from one another by an insulator (asbestos) layer 30 nm thick. For such a system the temperature dependence of resistance in the region of the superconducting transition has a temperature spread ΔT, which is due to fluctuations significant for such thin elements (width of the critical region). It is known that fluctuations smear the superconducting transition, so that the resistance smoothly decreases from the normal-state value R_N to zero in some finite temperature region. In the low temperature region, the fluctuations destroy the superconductivity and result in the appearance of finite conductivity. In the high-temperature part of the superconducting transition, an additional conductivity correction associated with fluctuation electron pairing appears. For mercury nanowires of 8-nm diameter the temperature smearing of the transition is $\Delta T \sim 0.4$ K (from ~4.9 K to ~5.3 K). Measurements performed on mercury nanowires with diameters of 2 nm [294] have shown that their superconducting transition is smeared from ~2 K to ~6 K, that is, $\Delta T \sim 4$ K. Some attempts to describe theoretically the fluctuation-smeared superconducting transition have been reported [476–480]; they are reviewed in [481]. The dimensionless smearing of the superconducting transition due to fluctuations in the 1D case (for ultrathin wires) was estimated to be

$$\Delta t_1 \equiv \Delta T/T_c = (\varepsilon_F/kT_c)[k_F^3 \xi(0) D^2]^{-1} \qquad (18)$$

where ε_F is the Fermi energy, $(h/2\pi)k_F$ is the Fermi momentum, $\xi(0) = (\xi_0 l)^{1/2}$, ξ_0 is the coherence length, l is the mean free path for electrons, D is the diameter, and T_c is the critical temperature. For macroscopic D (~0.01 cm), the dimensionless smearing can be calculated as $\Delta t \sim 10^{-12}$, and so no influence of fluctuations can be observed. For $D =$ 6 nm, Δt is about 10^{-1}, and for $D = 2$ nm, $\Delta t \sim 1$, which is in agreement with experimental data for mercury nanowires for $T_c \approx 4$ K. It is not quite clear how to determine the critical temperature T_c for such a temperature-smeared superconducting transition. This parameter was determined for superconducting nanowires within chrysotile asbestos channels from the temperature dependence of resistance, using theoretical descriptions of the fluctuation-smeared superconducting transition. In the low-temperature range, the temperature dependence of resistance can be written as

$$R(T) = R_N(\varepsilon/\Delta t_1)^{9/4} \exp[-(-\varepsilon/\Delta t_1)^{3/2}] \qquad (19)$$

where $\varepsilon = (T - T_c)/T_c$ is dimensionless temperature. T_c can be obtained by comparing the above dependence with experimental data. The temperature dependence $[\ln(R(T)/R_N)]^{-2/3}$ in the low-temperature part of the superconducting transition is linear and equal to zero at T_c. A good agreement with such a dependence was observed for all of the samples studied, and the T_c values were determined for different superconductors. The critical temperature T_c can also be calculated from the high-temperature part of the superconducting transition, where the temperature dependence of the resistance corresponds to the theoretical dependence

$$(\sigma - \sigma_N)/\sigma_N = \varepsilon^{-3/2} \qquad (20)$$

where $\sigma(T) \equiv 1/R(T)$ and $\sigma_N \equiv 1/R_N$. In this case, the dependence $[(\sigma - \sigma_N)/\sigma_N]^{-2/3}$ vs. T is linear and equal to zero at T_c. There was also a good agreement of such a dependence with experimental data, and T_c parameters determined for the two sides (low- and high-temperature) of superconducting transitions were the same. This is an interesting experimental result, because the theoretical descriptions of the low- and high-temperature parts of the superconducting transition (and, therefore, the experimentally obtained T_c parameters) are independent. Such a result indicates that this T_c corresponds to a physical quantity that can be designated the critical temperature for a smeared superconducting transition. Dependences of T_c on diameter D have been measured for mercury, tin, and indium nanowires with diameters of 2 to 15 nm [152, 299]. The critical temperature of tin nanowires increases with decreasing diameter (up to 5.2 K), and the $T_c(D)$ dependences for mercury and indium nanowires show maximums. For mercury nanowires, $T_c^{max} = 4.6$ K at $D = 4$ nm; for indium nanowires, $T_c^{max} = 6.5$ K at $D = 2.5$ nm.

The current-induced breakdown of superconductivity in nanowires within asbestos nanotubes has been studied [293, 294], and some discrete steps in voltage–current characteristics were observed. It was suggested that they correspond to nonequilibrium processes in superconducting nanowires, associated with formation of the so-called phase-slip centers (see [482, 483]), studied for superconducting whiskers and nanobridges. The voltage–current characteristics of mercury nanowires are strongly nonlinear in the superconducting region, and further cooling ($T < 2.9$ K) gives rise to a falling region [295]. It was found that the distinguishing features of the characteristics observed are the N-type shape and the existence of an upper limit to the oscillation frequency. The frequency dependence of the characteristics was investigated with a relaxation oscillator in which the sample served as the active element. The N-type region of the characteristics was accounted for by the appearance of a thermal resistive domain associated with local heating of the sample by the current.

The heat capacity of mercury and indium nanowires in the temperature range of the superconducting transition has been measured [296]. Experiments have been carried out on mercury nanowires with different diameters (from 2 up to 15 nm) and indium nanowires with a diameter of 5 nm, obtained by forcing the metal into the channels of chrysotile asbestos. A nonstationary method has been used (similar to that in [484, 485]) from the variation of the nanowire resistance in the region of the superconducting transition under the influence of a thermal flux modulating the temperature of the sample at a frequency of 25 Hz. This method fails to give the absolute values of the heat capacity and yields only its temperature dependence in the region of the superconducting transition smeared, because of fluctuations, over a considerable temperature interval. Three heat capacity peaks were observed in the region of superconducting transition for nanowires, and only one jump, in the bulk metal. This splitting of the heat capacity peak grows with decreasing nanowire diameter and vanishes when the nanowires are thick. These distinctive features of the heat capacity of metal nanowires cannot be accounted for by the coexistence of different crystal modifications with different critical temperatures (for mercury, α-Hg with $T_c = 4.15$ K and β-Hg with $T_c = 3.95$ K are known), because, first of all, there are three of them, and, furthermore, the temperature dependence of the heat capacity of indium nanowires, which do not have any different crystal phases, also displays three peaks. To explain the observed effect, it was suggested that the heat capacity peaks result from the electron system splitting into a series of subbands associated with spatial quantization and having their own different critical temperatures. In such a system, both the superconducting and the Peierls phase transitions (exhibiting different heat capacity features) are also possible [297, 486].

Critical magnetic fields have been measured for superconducting mercury nanowires [474, 475]. For nanowires 6 nm in diameter, the critical magnetic field extrapolated to $T = 0$ was 65 kOe for a magnetic field aligned with the nanowires and 36 kOe for a field oriented in the perpendicular direction. For mercury nanowires with a 2-nm diameter the critical magnetic field along the nanowires for $T = 0$ was 220 kOe, which exceeds by approximately a factor of 500 the value for bulk mercury. What is more important, this value is approximately three times the paramagnetic limit [457, 458] (77 kOe at $T = 0$). Such a large enhancement of the critical magnetic field can be explained in the framework of the theory, taking into account the spin-orbit scattering on the inner surface of asbestos channels [487].

4.2.2. Weak Localization in Nanowires

The electronic transport properties of thin wires in the temperature and field regime have been the object of numerous studies, mainly because of localization effects that strongly depend on the spatial dimensionality. In a 3D system, localization of all of the electronic states does not occur until the randomness exceeds a certain nonzero level. However, it has been shown for 1D systems [488, 489] that any amount of disorder, no matter how small, would make all of the states localized, and this would lead to vanishing conductivity at $T = 0$. For not strictly 1D systems (i.e., wires) it has been shown [490, 491] that the electronic states will be localized with a localization length L_{loc} equal to a wire length corresponding to an impurity resistance of $h/e^2 \approx 25.8$ kOhm. Thus, any wire with a resistance greater than this value, which is longer than the localization length, will have a thermally activated conductance at low temperatures and be an insulator at absolute zero. At high temperatures, an electron will experience frequent inelastic collisions with phonons. Each of these collisions will cause the electron to jump from one localized state to another, and the localization will have no effect on the conductance. The effect of localization will be felt when the mean free path for inelastic scattering is comparable to the localization length. At finite temperatures the resistance will be given by $R = R_0 + \Delta R$, where $\Delta R/R = L_i/L_{loc}$, R_0 is the temperature-independent impurity resistance, L_i is the distance to which an electron diffuses between inelastic collisions ($L_i = (D\tau_i)^{1/2}$), D is the electron diffusion constant (due to elastic collisions), and τ_i is the inelastic scattering time. Experimentally, the localization effects in thin wires have been studied in metals for samples fabricated with a method based on the substrate

step technique, in which a wire is formed in the corner of a step ion-milled into a substrate (see [492–495]). As for nanostructures within porous matrices, localization effects have been observed for mercury nanowires with diameters of 2 to 10 nm within the channels of chrysotile asbestos [496]. In the temperature range 10–50 K, the experimental data for Hg nanowires are well described by a $\Delta R \sim T^{-3/2}$ dependence (for 25-Å nanowires $\Delta R_{max}/R_0 \approx 0.4$). Below 6 K, the superconducting fluctuations govern the conductivity. As for nanowires with larger diameters, their negative temperature derivative of resistivity also corresponds to the $\Delta R \sim T^{-3/2}$ dependence, but $\Delta R/R_0$ decreases with increasing diameter. From the $\Delta R/R_0 \sim T^{-3/2}$ dependence it follows that $\tau_i \sim T^{-3}$. According to [490], such a temperature dependence of the inelastic scattering time may be due to scattering on 3D phonons. This is not surprising, since the nanowires are embedded in the dielectric matrix and are, to some extent, in contact with it, so that part of the phonons in such a composite medium may be three-dimensional. At 10 K, $\Delta R/R_0 \sim 0.45$, $\rho \sim 3 * 10^{-5}$ Ωcm (this value is determined from the ratio R_{300}/R_{10} and the resistivity of liquid mercury, $\sim 10^{-4}$ Ωcm), $D = v_F l/3 * 10$ cm^{2}s^{-1} ($v_F \sim 10^8$ cm/s, l is the mean free path for elastic collisions, determined from the temperature dependence of resistance). From these data, it can be found that $\tau_i \sim^{-10}$ s at 10 K. The localization length ($D\tau_i^{1/2}$) at 10 K is about $3 * 10^{-5}$ cm, which satisfies the localization criterion for these nanowires. More recently, localization effects have also been studied in Bi nanowires. The semimetal Bi is often selected for such studies because the electron density is about five orders of magnitude smaller than that in the conventional metals. Furthermore, the very small effective mass of electrons in Bi results in a large spatial extension of the electron wave functions; the effects of reduced dimensionality can, therefore, be seen in samples with dimensions on the order of 100 nm. Bi nanowires have been electrodeposited into nanometer-size cylindrical pores in polycarbonate membranes [497], introduced by pressure injection of a liquid Bi melt [498, 499] (see also [500]) and by vacuum evaporation [501] into the nanochannels of an anodic alumina template, and injected into porous Vycor glass by the application of hydrostatic pressures of 5 kbar [502]. The anodic alumina templates, having an array of parallel nearly cylindrical channels, were produced by anodizing aluminum substrates in acid solutions (for nanowires within these templates, see also [503–505]). The diameters of the pores in alumina templates are uniform to within 10% over the channel length (40–64 μm) from approximately 30 nm up to approximately 200 nm for different samples with average spacing between nanochannels of about 100 nm. In pores of polycarbonate membranes, Bi nanowires with diameters of 2000, 1000, 400, and 200 nm were fabricated, with wire densities of, respectively, $2 * 10^4$, $2 * 10^5$, $1 * 10^6$, and $3 * 10^6$ wires/mm^2. In the Vycor porous glass, the average pore diameter was about 6 nm. Measurements of the resistance of Bi nanowire arrays with different wire diameters have been carried out over a wide range of temperatures and magnetic fields. An increase in resistivity and a large positive magnetoresistance were observed. The additional magnetoresistance was ascribed [501] to a transition from 1D localization at a low field to 3D localization at a high field, when the magnetic length becomes smaller than the wire diameter. Most of the experimental results obtained in [499] are in good agreement with the theory and are accounted for by the electronic subband structure of quasi-1D Bi nanowires. Evidence that localization effects occur at low temperatures ($T < 4.0$ K) was obtained, but it was established that the localization effects are not the dominant mechanism affecting either the resistivity or the magnetoresistance in the temperature range 2.0 K $< T <$ 300 K. In Bi nanowires, 6 nm in diameter, in Vycor glass [502], no strong localization was observed, and this composite was a basically a good conductor. The observed temperature rise at low temperature and the associated magnetoresistance were interpreted in terms of weak localization.

4.2.3. Thermoelectricity in Bi Nanowires

Bi nanowires are of special interest for thermoelectric applications owing to the unique properties of bulk Bi, such as its small electron effective mass components, high anisotropy of its Fermi surface, and low thermal conductivity of Bi. Theoretical calculations predict that nanowires of bismuth must have an enhanced thermoelectric figure of merit [506] (see also [507, 508]), defined by $ZT = S^2\sigma T/k$, where S is the thermoelectric power (Seebeck coefficient), σ is the electrical conductivity, and k is the thermal conductivity. For a material to be a good thermoelectric cooler, it must have a high figure of merit, ZT. With known conventional solids, a limit for the figure of merit is obtained, and modification of any one of the parameters adversely affects the other transport coefficients, so that the resulting ZT does not vary significantly. Currently, the highest $ZT \cong 1$ at 300 K is observed in the Bi_2Te_3 compounds. It has been demonstrated that quantum wires have a strongly different ZT, since the electrons are confined to a single dimension. This increase is due mainly to the change in the density of states. In addition, there is enhanced phonon scattering from the wire surfaces. This must reduce the lattice thermal conductivity and, hence, lead to higher ZT. According to calculations, an array of nanowires with diameters on the order of 7 nm, oriented along the trigonal direction, could have a figure of merit of ~ 2 at 300 K. Single-crystal bismuth nanowires 200 nm in diameter (embedded in porous anodic alumina) have been used to measure the thermoelectric power and longitudinal magneto-Seebeck coefficient, and the first data on thermoelectric power, taken on an array of bismuth nanowires, were presented [509]. The temperature-dependent thermopower data are consistent with the partial electron and hole thermopower values calculated using the carrier Fermi energies obtained from Shubnikov-de Haas oscillations on the same samples.

4.2.4. Luttinger Liquid-Like Behavior in Semiconductor Nanowires

InSb nanowires with a diameter of about 5 nm (a length of about 0.1 cm) have been prepared within the channels of chrysotile asbestos [300, 302] by filling with InSb melt at 550 °C under a high pressure of 15 kbar. These nanowires are stable at room temperature, even after the pressure is removed. The temperature dependence of their zero-field electric conductance $G(T)$ is a power function of temperature, $G(T) \sim T^{\alpha}$, in the temperature interval 1.5–300 K

[302], with the exponent α ranging from $\approx$2 to $\approx$7. Current–voltage characteristics of such nanowires are nonlinear and follow at low temperatures the power law $I \sim V^{\beta}$ [302]. A similar behavior has been observed for 5-nm Te nanowires [301]. For InSb nanowires with 5-nm diameter and carrier effective mass $m^* = (0.014\text{–}0.2)m_e$ (possible range of effective masses for bulk InSb), the energy spacing between the first and the second quantum levels was $\Delta E = (\pi h)^2/2m*d^2 = 800\text{–}20000$ K $\gg T$ for the temperature range examined. So the electron band structure is in this case essentially one-dimensional. Thus, the samples studied consist of long quantum wires with one or few quantum conduction channels. The electron–electron correlation, negligible in the 3D case, predominates in the 1D case [510]. As a result, the physical properties of a 1D metal are expected to be dramatically different from the properties of the conventional metals with a Fermi liquid of electrons. One of the most significant consequences of the correlation effect is the absence of quasi-particle excitations in 1D metals. Instead, collective excitations associated with separate spin and charge degrees of freedom develop in the 1D case [511]. In the absence of long-range interactions, a 1D liquid (so-called Luttinger liquid) is formed [510, 511], whereas the long-range Coulomb interaction leads to a 1D Wigner crystal [512]. Charge transport is of a collective nature in this case and cannot be described by the conventional kinetic equation. The following temperature dependence of conductance, $G \sim T^{\alpha}$, has been predicted for tunneling between two drops of a pure Luttinger liquid, and nonlinear current–voltage characteristics, $I \sim V^{\beta}$ [513], are similar to those observed in InSb nanowires. It was suggested that, in multiple InSb samples prepared within asbestos nanotubes, long-range interactions between electrons in each wire may be screened by the Coulomb interaction of these electrons with electrons of neighboring wires. This leads to short-range intrawire electron–electron interaction, which is a basic assumption of the Luttinger liquid theory. Transport properties of individual nanowires are determined by impurities and weak links (e.g., constrictions) appearing in the fabrication process. The number of such weak links can be estimated, and it corresponds to $\approx 10^3$ weak links/cm per nanowire. The magnetoresistance of InSb nanowires has been studied over the temperature range 2.3–300 K in magnetic fields of up to 10 T [303]. The magnetic field leads to a 20% increase in the exponents α and β at $H = 10$ T. The magnetoresistance is positive; current suppression by up to an order of magnitude was observed at $T < 5$ K and $H = 10$ T. This may result from breaking of spin-charge separation in the 1D electron system, which is a novel mechanism of magnetoresistance.

The results of an experimental investigation of the thermoelectric properties of InSb and Te nanowires (diameter 5 nm) [301] support the Luttinger liquid model. The thermopower S of InSb nanowires is characterized below $T \approx 250$ K by a linear temperature dependence $S(T)$ ($G(T) \sim T^{2.8}$). For Te nanowires, $S(T)$ exhibits a metallic behavior but is not linear in any temperature interval and changes sign at $T \approx 180$ K ($G(T)$ for Te nanowires is $\sim T^{3.3}$). Such a behavior of the transport properties is typical of neither semiconductors nor metals and corresponds, at least qualitatively, to the Luttinger liquid model [514], which predicts a metallic behavior for the temperature dependence of thermopower (see also [515]).

4.3. Magnetic Properties

The influence of reduced physical dimensions on magnetic entities is of both fundamental and technological interest. Mainly motivated by technological interests, techniques for fabrication and characterization of magnetic nanoscaled systems have been developed. Physical vapor deposition techniques, followed by lithography of ever-increasing resolution, are usually considered to represent the ultimate technique for producing nanoscale magnets. However, well-crystallized nanomagnets can also be fabricated by a different method. In this method, thin wires of magnetic metals with large aspect ratios are electrochemically synthesized within the voids of anodic oxide films on aluminum [516] (Co and Co-Ni alloy) and track-etched polymer membranes [517] (Co, Ni). Nickel nanowire arrays have been fabricated by electrodeposition into the nanopores with lateral dimensions as small as 30 nm [518, 519]. The porous templates were fabricated by nuclear track etching. Particle tracks were formed in 5-μm-thick mica wafers by exposure to $\sim$6 meV α particles from a 100 μCi Cf-252 source in a chamber at a pressure of about 10^{-3} Torr. Interesting properties are expected to appear when the geometrical dimensions of the wires become comparable to a characteristic length scale, such as domain wall width or exchange length, and to mesoscopic dimensions, such as domain width. In bulk magnetic systems, the correlation length ξ increases with temperature and diverges at the bulk transition temperature $T_c(\infty)$. When one or more dimensions in the system are small, the growth of ξ is eventually limited by the smallest dimension d, and the system displays a lowered transition temperature $T_c(d)$ owing to finite-size effects. For Ni nanowires with diameters of 500 to 30 nm [519], the measured $T_c(d)$ values obey the finite-size scaling relation $[T_c(\infty) - T_c(d)]/T_c(\infty) = (\xi_0/d)\lambda$, where $\lambda = 0.94$ and $\xi_0 = 22$ Å. The measured correlation length $\xi_0 = 22$ Å is close to the value of 20 Å reported for thin polycrystalline nickel films. Large coercive fields have been observed for assemblies of ferromagnetic Ni cylinders with diameters ranging from 35 to 250 nm, produced by electrodeposition in nanoporous membranes [520]. At low temperature these coercive fields could be attributed to the curling mode of magnetization reversal, with account taken of the distribution of wire diameters and orientations. The coercive field of the nanowires of smaller diameter decreases from 1500 Oe to 200 Oe at 300 K nearly linearly. The magnetoresistance of similar nanowires has been measured at room temperature [521, 522]. The full magnetoresistive hysteresis loop was studied as a function of the angle between the applied field and the wire direction. The anisotropic magnetoresistance and magnetic properties of arrays constituted by similar objects have also been studied by other authors [523–526]. A giant magnetoresistance of about 15% has been observed at room temperature in a nanostructured material consisting of multilayer magnetic nanowires (Co/Cu) formed by electrodeposition into nanometer-sized pores of a template polymer membrane [527].

The classical antiferromagnet MnO has been embedded in a porous glass with an average pore diameter of about 7 nm [317]. MnO was synthesized from manganese nitrate solution by the chemical bath deposition method. The temperature evolution of the crystal and the magnetic structure of MnO under the "confined geometry" conditions were examined by means of neutron diffraction. It was demonstrated that a magnetic order similar to that observed in the bulk appears in regions smaller than the average size of nanoparticles. However, the ordered magnetic moment of 3.84 μ_B/ion was noticeably smaller than that in the bulk. The magnetic phase transition was found to be second-order with elevated Néel temperature, in contrast to the discontinuous first-order transition in the bulk. The experimentally observed temperature dependence of the magnetic moment was $m(T) \sim (1 - T/T_N)\beta$ with $T_N = 122$ K and $\beta = 0.34$. The critical index β is close to those expected for phase transitions in the classical 3D Heisenberg model.

4.4. Freezing and Melting in Confinement: Structure of Solid Phases

The confinement effect on freezing and melting (f/m) has been of considerable theoretical and experimental interest for many years. As early as 1888, J. J. Thomson proposed that the freezing temperature of a final particle would depend on the properties of the surface [528]. The first experiments on f/m in confined geometry (CG) were reported in the beginning of the last century [529–536].

Confinement usually results in depression of the f/m phase transition. Considering the difference in the Helmholtz free energy between liquid and solid in the pore, one obtains [537]

$$\Delta G = A[\gamma_{li} - \gamma_{si} + V\Delta G_m] \tag{21}$$

where ΔG_m is the free energy of fusion, γ_{li} is the substrate-liquid energy, and γ_{si} is the substrate-solid interfacial energy. A is the total interfacial area between the material and substrate and V is the total value of the confined material. The energies of any other interfaces in the system are ignored. ΔG_m is given by

$$\Delta G_m = \frac{[T_m - T]\Delta S_m}{V_m} = \frac{\Delta T \Delta H_f}{V_m T_m} \tag{22}$$

where ΔS_m is the entropy of melting, T is the melting point of the confined material, ΔH_f is the heat of fusion, and v_m is the molar volume. The differences in molar volume and heat capacity between solid and liquid are ignored.

Analyzing Eq. (21), one can determine two different characteristic temperatures [537, 538],

$$T_1 = T_m - \frac{A}{V}\frac{V_m T_m[\gamma_{si} - \gamma_{li}]}{\Delta H_f} \tag{23}$$

where A/V is geometry dependent and can be written as $A/V = \alpha/r$, where $\alpha = 3$ for a sphere of radius r and $\alpha = 2$ for a cylinder of radius r. T_1 corresponds to the equilibrium condition, when the energy of the completely liquid particle is equal to the energy of the completely solid particle. So T_1 represents [538] a thermodynamic lower bound for the melting temperature. However, taking into account kinetic arguments, one can obtain a second characteristic temperature equal in the case of a cylindrical pore of radius r

$$T_2 = T_m - \frac{3V_m T_m[\gamma_{si} - \gamma_{li}]}{r\Delta H_f} \tag{24}$$

T_2 corresponds to the vanishing of the energy barriers between solid and liquid. It represents a thermodynamic upper bound for the melting temperature. It is most likely that the freezing will take place around T_2, and the melting around T_1.

Following [537], we should emphasize that the influence of the material of the pore on the confined substance was completely neglected in the treatment considered above. The situation in the case of f/m phase transition is more complicated than that in the case of liquid/vapor. Wall effects on solid structure are expected to have a long-range character. The structure of the solid in confinement can be different from that the bulk.

In the following paragraphs we will present a brief summary of experimental studies of the f/m in CG (for a more detailed review see, e.g., [537] and references therein).

The first experimental results for f/m in confinement were mostly related to the f/m of water adsorbed on different hydrogels [529, 530]. The essential freezing-point depression was found for several different liquids on silica gel [536], and in the case of water adsorbed on 1.1-nm silica gel no freezing was found down to 208 K [539]. Scientific activity in the field essentially increased with the appearance of porous glasses, xerogels, and more recently channel matrices like MCM-41. f/m was studied for different types of materials.

Properties of gases and some of liquids embedded by absorption from vapors strongly depend on the degree of filling, f. At low filling the pore condensate exists as an adsorbate on the pore walls [227]; at higher filling the pore center is filled. So a distinction should be made between the layers absorbed on the pore walls and the fluid or solid in the central part of the pore.

4.4.1. Inert Gases

Inert gases (IGs) except helium (see Section 4.7) are the simplest substances to be studied in CG. In [540] suppression of the T_m for Ne in Vycor ($d = 5.4$ nm) (as well as for H_2 and D_2) was found, and estimation of the solid-liquid surface energy was found to be 2.6 mJm^{-2}. The 1.8 K difference between $T_m = 22.5$ K and $T_f = 20.7$ K (freezing temperature) was found and was attributed to the nucleation effects. In [228] f/m of neon and argon (together with H_2 and O_2) in Vycor and a silica xerogel were studied by combined heat-capacity and ultrasonic methods. Depression of the f/m was confirmed, but essential irreversibility of the freezing was found. It was demonstrated that any solid that forms, even in very small amounts, remains frozen until warmed well above the onset of solidification. Melting, however, is complete at a well-defined temperature between the onset of freezing and the bulk T_m. This fact was used as an indication that freezing is controlled by the pore geometry and not by nucleation kinetics. The width of the freezing region was attributed to the distribution of the pore sizes. Adams et al. [541] saw such measurements as providing a "spectrometer

of pore size." Wallacher and Knorr [542] presented results of elaborate heat capacity and vapor pressure study of Ar in Vicor. Measurements were performed for various filling and sample histories. It was demonstrated that first and second monolayers on the pore walls do not participate in f/m, whereas for other layers, in particular for the third one, a delayering transition was observed in the case of incomplete filling. Structures of frozen IG in CG were studied by X-ray diffraction [543, 544]. In the case of complete filling for pores down to 6 nm, a crystalline structure was revealed. Ar 2.2-nm and 2.5-nm diffraction patterns were similar to those of amorphous solids. For Ar in 7.5-nm gelsil in the case of incomplete filling diffraction patterns changed from amorphous at low $f \leq 0.4$ to crystalline at $f > 0.4$ [544]. The structure of the solid phase of Ar and Kr in Vicor [545] was interpreted as a disordered hexagonal close packed with phase transition to the fcc one at $T/T_f \approx 0.5$. In [544] this result was argued, and the structure of confined solid Ar was described in terms of fcc structure similar to that of the bulk, but with numerous stacking faults.

4.4.2. Diatomic Compounds: N_2, O_2, CO

In general f/m of diatomic gases in confinement is similar to the IG [228, 537]. In [546] a picosecond optical technique was used to study f/m of O_2 in Vicor. For low filling (less than two monolayers) no freezing was found. For pore sizes 2.2 nm $\leq d \leq$ 5.2 nm, ΔT_f followed the $1/d$ law (see Eq. (23)). f/m hysteresis was especially well pronounced at 2.2 nm and 2.8 nm. A γ–β solid-solid phase transition was observed, and the transition temperature $T_{\gamma\beta}$ was also depressed as $1/d$. A neutron diffraction study [547] demonstrated that the crystallites were considerably larger than the pores (about 10 times). An expected suppression of the melting temperature was also observed for nitrogen and CO. The crystal structure of CO is identical to structure of the bulk, with a crystallite size twice that of the mean pore diameter. In the case of N_2 and CO in 2.5-nm MCM-41 [537], it was suggested that the structure was either amorphous or liquid.

4.4.3. Metals

Unlike the cryogenic and H-bonded fluids, liquid metals usually do not wet the pores, and so there are no layering processes. For liquid mercury in the porous glasses, which has a high surface tension, it was demonstrated that the size of the solid clusters was temperature independent and equal to the pore diameter [276]. The main features of f/m transitions in Hg are broadening of both transitions and the existence of large thermal hysteresis, as demonstrated by neutron diffraction, NMR, and acoustic [274] measurements. It was shown that the freezing process was irreversible, but melting consisted of reversible and irreversible temperature regions. Combined use of longitudinal and transverse acoustic waves made it possible [274] to come to a conclusion about the origin of different behavior on melting. The broadening of melting was explained by the formation of a liquid layer on the mercury solid surface, whereas freezing was shown to be driven by the pore geometry with no visible precursor effects.

Few papers were published that related to the study of liquid metals in zeolites (Hg and Ga) [287], with cavities about 12 Å in diameter, and mordenite (Hg and Bi), with an average channel diameter of about 6.6 Å [548]. In 12-Å zeolites, well-defined jumps in temperature dependences of electric conductivity and peaks in the temperature dependences of the specific heat were observed and were considered to be evidence of the melting/freezing phase transition, with strong depression of the melting point. This result, which was rather surprising for such small clusters, was interpreted as evidence of the strong suppression of the fluctuations near the melting point related to the possible interaction between individual drops. In the case of the mordenite channel, the main result was demonstration of the possibility of producing monatomic metallic chains in the channels of the dielectric matrix; m/f transition was not searched for in this case.

Bogomolov et al. [549] have reported the study of one-dimensional metallic filaments produced by embedding of liquid mercury in the channels of natural chrysotile asbestos with diameters of 20–100 Å. It was demonstrated that as the filament diameter is reduced the phase transition region shifts toward low temperatures (shifts value follows the $\sim 1/d$ law) and becomes diffuse over a large temperature interval (the width of the transition region follows the $\sim 1/d^2$ law). The hysteresis between the melting and the solidification decreases and vanishes at $d \approx 20$ Å. Data analysis based on the Imry-Scalapino theory of the fluctuation effect on the first-order phase transition has demonstrated that the contribution of the fluctuations to the transition broadening does not exceed 1 K for the smallest diameters.

Heat capacity measurements of In in PG demonstrated essentially different behavior on melting and freezing (the heat capacity peak is much stronger for melting than for cooling). Careful differential scanning calorimetry (DSC) study of In in porous silica glasses with mean pore diameters from 6 to 141 nm has demonstrated [538] that, in agreement with Eq. (23), the melting temperature is reduced in inverse proportion to the pore size. In the smallest pores (8.2 and 18.2 nm) the latent heat of fusion was found to be about one-third to one-half of the bulk value. In was crystallized in the tetragonal phase identical to the bulk phase, but with a uniform lattice expansion of 0.5%. Unlike mercury, in the case of indium the crystallite size in 5.6-nm Vycor was about 20 nm, which is considerably larger than the pore size.

Liquid gallium in porous glasses and opals was extensively studied by electrical resistance measurements [550], X-ray diffraction [275, 550, 551], NMR [552–554], and acoustical [555, 556] methods. In contrast with mercury in porous glass, the size of confined gallium crystallites was estimated to be 22 nm, which was significantly larger than the pore size ($\approx$4 nm) [275]. Within the limits of experimental accuracy this size was nearly temperature independent. Only a slight tendency to the peak narrowing (i.e., size increase) was observed on cooling. The authors concluded that since the size of crystallites was about constant during the melting, the broadening cannot be attributed to the particle size distribution. Nor can it be explained in terms of the pore size distribution, since the gallium crystallites were much larger than the pores. As in the case of mercury, it was proposed that the broadening is related to the formation of the

liquid around the particles. Information about the crystal structure of the solid gallium confined in the porous glasses is contradictory. In paper [551] several modification were found upon Ga solidification in the 4-nm porous glass, while in [275] only a single phase was found down to 20 K, and below this temperature the appearance of the second phase was reported. Both modifications reported in [275] were different from the known bulk gallium structures. In the case of Ga confined in opal, four different modifications were reported [550]. Two of them were similar to those observed in the porous glass and so were different from bulk, whereas two others were identified as β-Ga and disordered α-Ga. An interesting effect was reported for the f/m of disordered α-Ga. Melting-freezing hysteresis for this modification was found to depend on the temperature of prewarming.

4.4.4. Water

As mentioned above, water was the first material used to study the f/m transition in a confined geometry. A number of methods were used for its study, including DSC [557], neutron diffraction [557–559], and NMR [560, 561]. The studies of water freezing in the porous materials (silica gel, porous glasses, and activated charcoal) showed a freezing point depression consistent with two or three layers of nonfreezing water on the pore surfaces [537, 560]. In some cases freezing resulted in the formation of confined crystallites with a structure different from that of the bulk. In particular, cubic ice was shown to form in pores at 260 K. This modification was known before only at high pressure conditions. In samples with 2-nm pores, water remained liquid down to 251 K. On lowering of the pore size from 50 nm to 4 nm, the predominant ice modification changed from hexagonal to cubic [562]. NMR studies of water confined in glasses with different hydrophobicities (treated with hexamethyldisilazane) have demonstrated an increase in the melting point depression with increase of hydrophobicity. In the case of the most hydrophobic samples water was unable to penetrate the pores. Such behavior is consistent with the larger contact angle of the water-ice interface on the treated glass substrate [537]. In [562] the low-temperature (down to 173 K) structure of water in MCM-41 with $d = 3.3$- and 3.5-nm structure was found to be incompatible with neither cubic nor hexagonal ice. X-ray diffraction experiments performed on water confined in 2.4 nm and 4.2 nm MCM-41 demonstrated that water freezes abruptly in the middle of 4.2-nm channels and more gradually in 2.4-nm channels [563]. As in the porous glasses, a disordered liquid-like layer was found. In contrast to porous glasses, no hysteresis was found in MCM-41 materials. This results in the conclusion that large f/m hysteresis in other matrices can be attributed to a network effect [537].

4.4.5. Organic Liquids

For organic liquids confined in porous glasses, depression of the freezing temperature is typically observed. For cyclohexane in silica glass the temperatures of both f/m transition and the transition from the plastic to the brittle crystal are decreasing [564, 565]. Booth and Strange [566] demonstrated that for pores smaller than 5 nm, the plastic crystalline phase was replaced with a structurally disordered phase with high rates of diffusion. No freezing was observed in the case of 2,4,6-trinitrotoluene in 2.5-nm and 5-nm gelsil glasses and in 10-nm and 20-nm glasses a bulklike orthorhombic structure was formed [567]. Systematic studies of CCl_4 and nitrobenzene in silica glasses and silica gel (4–100 nm) [537, 568, 569] showed a good linear correlation between the melting-point depression and the inverse pore diameter. Takei et al. [569] found no changes in the transition temperature resulting from silanizing the pores with hexadimethylsilazane. The same linear dependence was observed in [570] for *cis*-decalin, *trans*-decalin, cyclohexane, benzene, chlorobenzene, naphthalene, and *n*-heptane in silanized silica glasses (4–73 nm). No freezing was observed for cyclohexane or *cis*-decalin in 4-nm pores [537, 570].

Essentially different behavior was observed in the case of f/m of organic liquids in MCM-41 materials. Morineau et al. reported differential scanning calorimetry measurements of cyclohexane, benzene, toluene, *o*-terphenyl, and *m*-toluidine confined in 4-nm MCM-41 [571]. Only a moderate decrease (not exceeding 6 K) of the f/m temperature was observed. In the presence of a few percent excess liquid, crystallization of the confined part was found to be sensitive to the presence of crystal outside the pores. The absence of supercooling of the confined part was explained by a freezing mechanism induced by the outside part. Glass transitions in confined toluene, *o*-terphenyl, and *m*-toluidine were nearly unaffected by confinement.

4.5. Dynamics

Confinement strongly affects dynamical properties of the embedded materials, especially liquids (for the solid phase the problem is much less studied). Especially well studied are water and H-bonded systems [572, 573]. For many H-bonded fluids crystallization is easily suppressed in confinement, and these systems are considered as glass-forming materials. In general there are several mechanisms that can influence the molecular dynamics:

- Structural effects due to (i) geometrical obstacles in the finite volume [574] and (ii) a reduced density resulting from the difficulty in packing molecules. In the case of diffusion the existence of dead-ended pores became important.
- Surface effects: Interaction of liquids with the surface results in the formation of two (interfacial and bulklike) [575] or even three [576] (consisting of molecules having solidlike, interfacial, and bulklike dynamics, respectively) layers.
- Finite-size effects due to spatial heterogeneity of the confined material–when the length scale of confinement becomes smaller than an intrinsic length scale related to the dynamics of bulk material. Such a situation is typical in glass formers when the glass transition temperature, is approached.

Dynamics of confined materials are studied by several methods: dielectric spectroscopy [574, 575, 577–579], nuclear magnetic resonance (NMR), optical spectroscopy (Raman and dynamic light scattering) [573], and neutron spectroscopy [573, 579] are most important.

4.5.1. Dielectric Properties in Confinement

Dielectric properties of confined systems are crucial for many applications. Dielectric spectroscopy is appropriate for the investigation of polar materials and covers a large dynamical range from dc to microwaves [574, 580]. The difficulty with the interpretation is that neither the permittivity of the filler, ε_p, nor that of the matrix, ε_m, but a volume-averaged value, ε_{eff}, is obtained. ε_{eff} can be presented as [574]

$$\varepsilon_{eff} = \frac{(1-f)\varepsilon_m + f\varepsilon_p \langle E\rangle_p / \langle E\rangle_m}{(1-f) + f\langle E\rangle_p / \langle E\rangle_m} \tag{25}$$

where $f = V_p/V_{total}$ is the volume filling factor, $\langle E\rangle_p$ is the electric field averaged over the filler volume, and $\langle E\rangle_m$ is the electric field averaged over the matrix volume. The ratio $\langle E\rangle_p/\langle E\rangle_m$ and thus the measured effective permittivity depend on the microstructure, that is, the topology, degree of order, and dimensionality. This problem is discussed in detail in [574]. In general we can formulate the following main results:

- 3D confinement (droplets): The values of ε_{eff} and the relaxation frequency ν_{relax} depend on the exact spatial distribution of the particles. In the case of dilute inclusions of strongly polar materials, relaxation strength $\Delta\varepsilon_{eff} \ll \Delta\epsilon_p$ becomes independent of $\Delta\epsilon_p$. With increasing disorder or filling factor F, $\Delta\varepsilon_{eff}$ increases, while the frequency shift of ν_{eff} decreases. Agglomeration gives rise to a broadening of the relaxation peak [581]. In the case of asymmetric relaxations described by the Havrilak-Negami (HN) formula,

$$\varepsilon_{HN} = \varepsilon_\infty + \frac{\Delta\varepsilon}{\{1 + [i(\nu/\nu_0)]^{1-\alpha}\}^\gamma} \tag{26}$$

  with shape parameters α and γ, the shift of the relaxation frequency increases with increasing asymmetry ($\gamma < 1$). Similar to the symmetric case, it decreases with increasing disorder or filling factor. The slopes of the low- and high-frequency sides of $\varepsilon''(\nu)$ are unchanged.
- 2D confinement (channel-like pores running through the whole sample; real porous systems with a high amount of dead ends and/or a fractally rough surface should be considered as an intermediate 2D-3D case): In the case of noncrossing pores, $\nu_{eff} \simeq \nu_p$; $\Delta\varepsilon_{eff} \propto \Delta\varepsilon_p$). The high-frequency side of $\varepsilon''(\nu)$ is slightly flattened. In the case of a network of interconnected pores the form of electric relaxation is hardly distinguishable from that of bulk material. In the most important case of real porous systems, the following basic features can be formulated [574]:
  - The effective relaxation strength increases linearly with $\Delta\varepsilon_p$ for large enough values: $\Delta\varepsilon_{eff} \simeq a\Delta\varepsilon_p + b$.
  - The effective relaxation frequency roughly equals that of the filler, $\nu_{eff}/\nu_p \simeq 1$.
  - The shape of the relaxation is preserved on the low-frequency side of the peak and is weakly modified on the high-frequency side.
- 1D confinement (films or multilayer structures): The situation depends on the orientation of the layer. In the case of layers oriented parallel to the film, the characteristics resemble those of a 2D confinement, and in the perpendicular case, the 3D confinement.

The molecular dynamics of confined liquids in the frequency range from 10^{-6} to 10^{12} Hz is characterized by the superposition of different relaxation processes that take place on a local scale (β-relaxation) and cooperative fluctuations (α-relaxation, dynamic glass transition). In most cases an additional relaxation peak appears, which is usually ascribed to the Maxwell-Wagner-Sillars (MWS) polarization appearing when a slightly conducting liquid is enclosed in an insulating material [577, 582]. Experimental dielectric results were obtained for several different confined systems. In nitrobenzene in CPG, Vycor, and MCM-41 [582], already mentioned above, two relaxation components were found in the liquid state at about $5 * 10^{-3}$ s and 10^{-4} s, both of simple Debye relaxator shape. The slow one has been attributed to the MWS polarization, and the fast to the relaxation of the contact layer. The response of the bulk fluid was out of the experimental range. Nitrobenzene is not a glass former, and neither component demonstrates essential temperature dependence. The dynamic behavior of glass formers is essentially different. Several glass-forming liquids were studied, among them *N*-methyl-ε-caprolactam (NMEC) (nonassociating liquid) [577, 583, 584], propylene glycol (PG) and its oligomers (PPG) [578, 579, 584], salol [575], pentylene glycol [575], and glycerol [575]. Experimental dielectric spectra typically contained a MW peak plus one or two relaxation peaks. Peaks are broadened as compared with a simple Debye relaxator and asymmetric; the shape is consistent with the HN formula. In the case of two relaxation peaks a slow process was unambiguously attributed to the interface dynamics. This fact was confirmed by modifying the porous glass surface (silanizing with trimethylchlorosilane) [577], which results in weakening of liquid (NMEC) to the surface and in the nearly complete suppression of the slow peak. Liquids with two (PG) or three (glycerol) hydroxy groups were shown to exhibit only one relaxation process [575]. This was interpreted in terms of a shell model of a bulklike phase and an interfacial layer including molecular exchange between the two subsystems. For all components relaxation times were temperature dependent [575, 577, 579], following the Vogel-Fulcher-Tammann (VFT) relation [585, 586]

$$\tau_m = \omega_m^{-1} = A\exp[DT_0/(T - T_0)] \tag{27}$$

where ω_m is the position of the maximum of the dielectric loss peak, A is the prefactor, D is the fragility parameter (higher values of D correspond to stronger glass formers), and T_0 is the Vogel temperature. A fit of dielectric data to the VFT equation is often used to determine the glass transition temperature T_g, based on $\tau_m(T_g) = 100$ s. In such a definition T_g is exactly dynamic, determined by the whole set of VFT variables. Some published results are summarized in Table 7. For interface components relaxation is always retarded; as clearly seen from a comparison of prefactors, T_g for the interface layer is essentially lower than that for bulk. However, there is no regular change of Vogel temperature, even for the interface layer. The α relaxation of the liquids

Table 7. Parameters of VFT equation for dielectric relaxation.

System	A (ps)	D	T_0 (K)	T_g (K)	ΔT_g (K)	ΔT_0 (K)
PG(bulk) [579]	195	18.4	112.3	166.8		
PPG-400(bulk) [579]	9	7.8	160.4	196.4		
PPG-725(bulk) [579]	32	8.3	160.3	197.1		
PPG-4000(bulk) [579]	12.3	7.0	164.6	197.4		
PG(conf., inner) [579]	892	19.1	115.4	171.3	+4.5	+3.1
PPG-400(conf., inner) [579]	84	9.3	158.0	197.7	+1.3	−2.4
PPG-725(conf., inner) [579]	49	9.0	157.8	197.3	+0.2	−2.5
PPG-4000(conf., inner) [579]	11.6	6.9	166.1	199.0	+1.6	+1.7
PG(conf., interface) [579]	0.2	11.6	114	188	+21.2	+1.7
PPG-400(conf., interface) [579]	0.7	9.8	155	215	+18.6	−5.4
PPG-725(conf., interface) [579]	0.6	12.7	147	221	+23.9	−13.3
PPG-4000(conf., interface) [579]	$1.1 * 10^{-6}$	6.6	174	244	+46.6	−9.6
NMEC(bulk) [583]	0.23	3.38	142.3	173.5		
NMEC(confined) [583]	0.062	3.74	130.6	161.1	−12.4	−11.7
Salol(bulk) [575]	0.1	4.8	194	222		
Salol(confined 7.5 nm) [575]	$1.8 * 10^{-3}$	8.0	177	215	−7	−17
Salol(confined 5 nm) [575]	0.17	5.3	185	214	−8	−9

Note: $\Delta T_g = T_g^{conf} - T_g^{bulk}$, $\Delta T_0 = T_0^{conf} - T_0^{bulk}$.

in the inner pore space is much less affected by confinement and the results of different authors are controversial, as one can see from the data in Table 7.

Conducting Liquids in Porous Media In many natural inhomogeneous materials, such as brine-saturated sedimentary rocks, giant values (about 10^6) of low-frequency dielectric permittivity were observed ("dielectric anomaly of rocks"). A similar result was obtained for brine-saturated porous alumina ceramics [587], where $Re\varepsilon \approx 2 * 10^5$ was reported for the porosity 0.193. This anomaly was analyzed [587–589], and the best agreement was obtained for the model considering a system of pores separated by thin walls.

In [590] giant growth of ε was observed on heating of $NaNO_2$ ferroelectric confined in 7-nm porous glass (up to 10^8 at 100 Hz and $T > 550$ K). This effect was explained by an essentially broadened melting transition. In a very recent paper [591] the temperature evolution of structure in a restricted geometry was studied for the ferroelectric $NaNO_2$ embedded in a porous glass, and it was shown that this CM forms a kind of interconnected cluster, probably of the dendrite type, with a practically temperature-independent average size of about 45 nm. Above T_c the volume "premelted" state is formed, manifesting itself in a sharp growth of the thermal motion parameters, softening of lattice, and increase of lattice volume. In such a case the possible appearance of ionic current due to oxygen jumping diffusion is proposed as a reason for the observed giant growth of dielectric permittivity. On cooling below T_c macroscopic polarization and potential barriers suppress the lattice softening, and the normal ferroelectric phase exists.

There is very limited information on the behavior of *ferroelectric materials* in confinement. In [263, 265, 592] sodium nitrite and Rochelle salt embedded in 7-nm porous glass were studied. Only a very small negative shift of transition temperature was found in contrast to the dispersed $BaTiO_3$ and $PbTiO_3$ [593]. Substantial broadening of the transition was reported and interpreted as an effect of the fluctuations. In [263, 592] KDP embedded in artificial opal and porous glass was studied. A fast increase of the transition temperature with decrease of pores size was observed. The most plausible explanation is coupling of the embedded materials with the pore surface.

4.5.2. Diffusion

Hindered transport in restricted geometries is of theoretical and practical interest because of its relevance to many important processes, such as chromatographic separation of polymers, enhanced oil recovery, membrane separation, and polymerization in the presence of heterogeneous catalysts [594].

In an early paper on the study of molecular diffusion by forced Rayleigh scattering in porous media, published more than 15 years ago [595], the diffusion of an azobenzene molecule in Vycor was reported to be two orders of magnitude slower than in free unbound solution. The authors succeeded in separating chemical heterogeneity and geometrical disorder and developed a detailed fractal model relating pore size distribution to self-diffusion. In another paper [596] the diffusion of polystyrene in 74.5-nm porous silica glass was studied by dynamic light scattering. Adsorption of polymer was excluded by silanizing the surface. Decay of long-wavelength fluctuations was governed by a single "macroscopic" relaxation rate. It was found that the correlation length depends not only on the pore geometry but on the hydrodynamic radius R_H of the polystyrene molecule as well. Samples with number-averaged molecular weights of 35,000 and 93,000 corresponded to $R_H \approx 4$ nm and $R_H \approx 7.5$ nm, respectively. The following description was proposed for an effective macroscopic self-diffusion coefficient D_∞

$$D_\infty = \chi f(R_H/R_P) D_0 = \chi D_P \tag{28}$$

where χ is the intrinsic conductivity (hindrance factor) of the porous medium ($T = 1/\chi$ is called tortuosity), D_0 is the bulk diffusion coefficient, R_P is the average pore radius, and D_P is the diffusion coefficient within the pores for the

short times or distances. An intrapore short-distance translational diffusion coefficient can be effectively measured with incoherent quasi-elastic neutron scattering [597], though we do not know any paper in which coefficient f was analyzed on the basis of a comparison of D_∞ with D_P from neutron data. A rough estimate of $\chi \approx 0.75$ was given in [596] with estimations for f of 0.92 and 0.83, and 4 nm and 7.5 nm for R_H, respectively. Diffusion of polystyrene with molecular weights of 2500–13,000 in 2-nm Vycor glass with a porosity $\phi \approx 0.28$ (ratio of polymer free-solution radius to the glass pore radius from 0.017 to 1.4) was studied with dynamic light scattering [594]. Strong dependence of the D_∞/D_0 on R_H/R_P was confirmed, and an estimation of $\chi = 0.3$ to 0.7 was obtained. Such an estimation is in good agreement with calculations in the frames of model of medium composed of random structural elements (pores) chaotically connected with one another [598]

$$\chi = 1 - \tfrac{2}{3}(1+\phi)(1-\phi)^{3/2} \tag{29}$$

This yields $\chi = 0.48$ for $\phi = 0.28$. The extremely small χ value obtained in [595] was explained in [594] by incomplete suppression of the adsorption on the pore walls.

A limiting case of diffusion in the very narrow pores was considered in [599]. Diffusion of tetrafluoromethane (diameter of molecules 0.47 nm) in the $AlPO_4$-5 zeolite (channel diameter on the order of 0.73 nm) was studied by pulsed field gradient NMR. $AlPO_4$-5 zeolite is traversed by parallel channels. It was demonstrated that in this case the mean square displacement $\langle \zeta^2 \rangle$ increases in proportion with the square root of observation time, in contrast to the ordinary diffusion when $\langle \zeta^2 \rangle \propto t$. This observation corresponds to the regime of "single-file" diffusion, when the molecules are unable to pass each other. Measurements provide a diffusion coefficient D on the order of $5 * 10^{-7}$ $m^2 s^{-1}$, which is about 2 orders of magnitude larger than in other zeolites [600]. The observed phenomenon was explained in [599] as the effect of molecular guidance by the channels, resulting in the enhancement of molecular mobility in one-dimensional channels in comparison with higher-dimensional pore networks.

Self-diffusion properties of plastic formers succionitrile and cyclohexane confined in porous glasses [601, 602] and in MCM-41 [603] were studied. In the confined liquid, above freezing the diffusion coefficient is essentially lower than in the bulk, similar to the other confined substances considered above. However, on freezing, the diffusion coefficient in bulk drops down by about 4 orders of magnitude. In the confined material two phases develop, namely, a plastic phase in the center of the pore and a liquid-like component with D about 3 orders of magnitude higher than in the bulk.

4.6. Liquid Crystals

Liquid crystals (LCs) are often considered as ideal objects for the study the effects of confined geometry. Kralj et al. [249] have explicitly formulated the main advantages of LCs as model systems for such research: (i) LCs exhibit a variety of phases with different degrees of orientational or translational order; (ii) there exist different kinds of transitions between these phases; (iii) they are typical representatives of soft materials (their response to perturbations induced by the confining matrix is pronounced and long-ranged); (iv) LCs and the host materials do not interact chemically; and (v) both LCs and the host matrixes are in most cases transparent, and consequently samples can be studied by a variety of optical methods. Key to understanding the large part of confinement effects on LCs is an interaction of the LCs with surfaces [604]. In the case of the nematic LC this interaction is usually described in terms of the preferred orientation of the nematic director at the surface called the "easy direction" and the strength of the interaction. The easy direction is the direction of the director at which the surface energy is minimal. The strength of the interaction is determined by the so-called anchoring strength, which tells what torque is needed to move the orientation of the director from the easy direction. Anchoring properties of the surface strongly affect both static and dynamic characteristics of LCs.

The influence of the confinement on the LCs is related to the contribution of the LC/matrix surface to the free energy. This effect is usually described in the framework of the Landau-de Gennes theory [605, 606], with the surface term described as

$$F_s = \int (f_n + f_s) d^2 \vec{r} \tag{30}$$

with the nematic contribution f_n described as [607]

$$f_n = -W_1 S\left(\frac{3(\vec{n}\,\vec{e}^{\,2} - 1)}{2}\right) + W_2 S^2$$

where W_1 and W_2 are positive surface anchoring constants, $\vec{n}$ determines the nematic director field, S is the nematic orientational order parameter, and $\vec{e}$ is the easy direction. The linear term is concerned with the direct interaction between the surface and LC molecules, and the S^2 term is due to the fact that LC molecules have fewer neighbors at the surface [249]. The smectic contribution f_s is presented as [249]

$$f_s = -W_p \eta \cos(\Phi - \Phi_s)$$

where W_p is a positive constant.

We will follow [249] in describing possible LCs structure in the simplest case of the cylindrical cavities of radius R and length L_d or the cavities that can be presented as a space between two concentric cylinders of radii R_0 and $R_0 + 2R$, respectively. Two extreme cases can be considered: homeotropic anchoring (the LC molecules tend to be aligned along the surface normal) and planar anchoring (the molecules are forced to lie in the limiting plane, where all directors are equivalent). The case of planar anchoring corresponds to the homogeneous structure (Fig. 6(i)) or the homogeneous structure with point defects (Fig. 6(ii)). In the case of homeotropic ordering one expects formation of the escaped radial (Fig. 6(iii)) [249, 608] or escaped radial with point defects (Fig. 6(iv)) structures. In the limiting case of strong homeotropic anchoring, structures (iii) and (iv) are transformed into the planar radial structure (Fig. 6(v)) [249, 608, 609]. In this case the bend nematic deformation, characteristic for structures (v) and (iv), is absent and the director field radially streams from the center of the cylinder with a line defect along the cylinder axis.

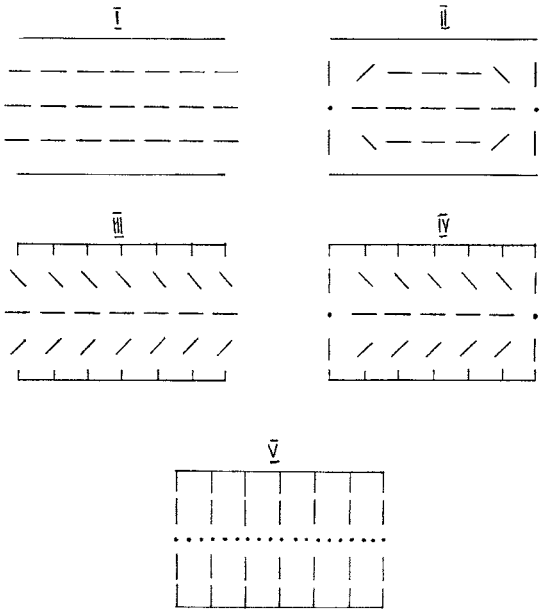


Figure 6. Model director structures. (i) Homogeneous. (ii) Homogeneous with point defects. (iii) Escaped radial. (iv) Escaped radial with point defects. (v) Planar radial structure. For the type (v) cavity the director field is either radially distributed or perpendicular to $\vec{R}$. (following [249]).

An important effect is related to the quenched disorder in the LC-matrix nanocomposites. The matrix in this case not only geometrically confines the LCs but also induces a random orienting field that fixes the direction of the order parameter near the surface of the matrix [610]. The randomness is determined by the geometry of a confining porous matrix. Maritan et al. [611] were the first to associate the effects of the porous media with that of a random field. Their theoretical considerations supported the early experimental results of [612, 613] reporting the light scattering and calorimetry measurements of confined LCs. It was demonstrated that the first-order phase transition observed in bulk LCs was replaced by a smooth evolution to a glassy state, with the correlation length not exceeding the characteristic pore size. More elaborate (renormalization group) theoretical treatment of the problem of confined nematics was recently reported by Feldman [610]. He has demonstrated that in the case of weak disorder a state with quasi-long-range order can be formed. However, in the case of strong disorder the topological defects drive the system into the glass state, in which the orientation of the director is determined only by the local random potential. Influence of randomness to a large extent is governed by the ratio between the average pore radius R and the nematic correlation length ξ_n. In [249] estimations were made for the controlled porous glass matrix, and it was shown that the effects of randomness become important for $R < \xi_n$, where the noncorrelated curvatures of cylindrical voids are the main origin of randomness.

Properties of the confined LCs have been extensively studied by different experimental techniques such as NMR [249, 614], light scattering [604], dielectric spectroscopy [615], X-ray diffraction [246], etc.

In particular it was shown that in small pores, as already mentioned, isotropic-nematic (I-N) phase transition is replaced by a gradual evolution of the nematic order [249, 612–614]. However, in the case of large enough pores ($R \geq$ 25 nm for pentylcyanobiphenil (5CB) in controlled porous glass [249]) the I-N phase transition is discontinuous. In this case $T_{\text{I-N}}$ can be shifted either upward or downward, depending on the size of the cavity and on the ordering properties of the cavity surface. In the mentioned example of 5CB in CPG, with a decrease of $T_{\text{I-N}}$ with $\Delta T_{\text{I-N}} = \mathcal{T}_{\text{I-N}}^{\text{bulk}} - \mathcal{T}_{\text{I-N}}^{\text{conf.}} \propto 1/R^{1.3\pm0.3}$ [249]. $\Delta T_{\text{I-N}} = \propto 1/L$ (L is the mean free path in the empty regions), scaling was observed for LCs in aerogel matrix [616]. It should be mentioned that in many cases the surface-induced paranematic order is observed in the isotropic phase [614]. The nematic-smectic-A phase transition is usually more affected than the I-N one [249]. It was demonstrated for the octylcyanobiphenyl (8CB) in controlled porous glass [246, 249] that 8CB was confined to anopores (200 nm in diameter) [617] and aerogels (mass densities $\rho = 0.080.60$ g $\cdot$ cm^{-3}) [616]. In the case of porous glasses, LC-surface interaction was modified by silanizing the surface. According to [249], silane-treated surface exhibits homeotropic anchoring and nontreated surface planar anchoring. In nontreated samples the smectic order parameter temperature dependence is similar to that in bulk LCs, with the order parameter reduced by a few percent [246]. In the case of nontreated surface pretranslation, smectic ordering is observed. Temperature dependence of the smectic correlation length ξ_{sm}, similar to that in magnetic systems with random fields, was observed for both treated and untreated samples [246]. ξ_{sm} was gradually increased by lowering the temperature from the nematic phase and saturated at $\xi_{\text{sm}} \approx R$. In the case of anopores the surface was either untreated or lecithin-treated. The rough untreated surface inhibited the smectic growth, whereas in the case of the lecithin-treated surface the smectic layers were initiated at the surface [617].

Dynamics of LCs is strongly affected by confinement. Broadband dielectric spectroscopy of 5CB and 8CB in porous glasses (10–100 nm) has revealed a number of differences between bulk and confined behavior [618]. Similar to those of confined liquids (see 4.5.1), all observed relaxation processes are of the non-Debye type but can be well described by the Havriliak-Negami equation, (26). The relaxation processes in confined LCs are not frozen, even at temperatures about 20° below bulk crystallization temperatures. Confinement resulted in the appearance of two new relaxation processes not observed in the bulk LCs. One of the processes is related to the presence of the surface layer; the origin of the other is not clear yet. Dynamics of the nematic-isotropic phase transition was studied with time-resolved transient grating optical Kerr effect measurements [619]. Nonexponential decay and drastic decrease in the relaxation time of the order parameter fluctuations were observed in the pretransitional temperature range. This result is in contrast with data of Wu et al. [612], who observed below bulk $T_{\text{I/N}}$ slowing down of the critical fluctuations described by the Vogel-Fulcher law.

Dielectric spectroscopy was also used to study the smectic-A to smectic-C* transition in W314 LCs confined in aerogels ($\rho = 0.08$ and 0.17 g $\cdot$ cm^{-3}) [615]. In addition to

the Goldstone collective mode, an additional relaxation process related to the surface layer was found with decreasing, strength while the strength of the Goldstone mode increased on cooling.

4.7. Helium in Confinement

The problem of quantum liquids (He and hydrogen) is too broad to be covered as a small part of a more general review, so we will indicate just some of the main topics related to this problem. As early as 1956, the first studies of Vycor glass as a superleak were reported in [620]. Later on, solidification [537, 541], superfluid transition [621–624], dynamics [625, 626], and other properties were studied in great detail. A paper by Adams et al. [541] remains the most detailed study of the P-T phase diagram of ^{4}He confined in the Vycor glass for 0.8 K $< T <$ 2.5 K and $P \leq 60$ bar. No solidification was found for $P < 36.3$ bar, and it was shown that at $T = 0.8$ K, solid and liquid may coexist for 36.3 bar $< P <$ 38.0 bar. The line of λ superfluid phase transitions was shifted to lower temperatures for about 0.2 K but remained parallel to that in bulk liquid He. In [623] the temperature dependence of the superfluid density ρ_S was examined in the critical region for ^{4}He in Vycor glass. It was shown that superfluid density follows a power law in reduced temperature, $\rho_S \propto \tau^{\zeta}$, with the exponent $\zeta \approx 0.67$, in numerical agreement with the bulk value. Later [624] $\rho_S(\tau)$ dependences were checked for He in Vycor (full filling and films) xerogel (full and film) and aerogel, and values of ζ from 0.63 to 0.91 were obtained. Different critical exponents indicate different universality classes. Moreover, Mulders et al. [627], analyzing earlier published data on ρ_S and specific heat temperature dependences, have demonstrated apparent violation of the hyperscaling relation $3\zeta = 2 - \alpha$. We are not aware of any reliable explanation of the described facts. Recently several papers were published that attempted to solve the puzzle of confined helium properties by microscopic measurements. In [626] Sokol et al. report the results on an inelastic neutron scattering study of roton excitations in two different porosity aerogels. Crossover in the temperature dependence of the roton energy gap was observed at $T \approx 1.9$ K, which was explained in terms of length scales of confinement and the roton mean free path. It was suggested that the roton mean free path determines the spatial scale above the crossover temperature, while below it the temperature dependence of the roton energy is suppressed. Some very recent inelastic neutron scattering results on He in xerogel glass were presented in [625]. In spite of the huge total number of publications, we believe that there is still no comprehensive understanding of the problem.

GLOSSARY

Confinement Restriction of some material to a given volume; can be one-dimensional (films), two-dimensional (filaments), or three-dimensional (isolated particles).

Phase transition Change of a substance from one phase (the type of a system, such as solid, liquid, gas, superconductor, ferromagnet, etc.) to another. Also known as phase transformation.

Porosity The existence of some free volume inside a solid material structure, which is not occupied by structural elements of this material.

Size effect The effect of size of a piece of material on its properties.

Wettability The ability of any solid surface to be wetted when in contact with a liquid; that is, the surface tension of the liquid is reduced so that the liquid spreads over the surface.

REFERENCES

1. P. G. Cheremskoy, V. V. Slezov, and V. I. Betehtin, "Pores in Solid State." Energoatomizdat, Moscow, 1990.
2. R. Hilfer, *Phys. Rev. B* 44, 60 (1991).
3. F. Rodriguez-Reinozo, "Characterization of Porous Solids." Elsevier, New York, 1991.
4. T. G. Plachenov and S. D. Kolosentsev, "Porosimetry." Chemistry, Leningrad, 1988.
5. A. J. Katz and A. H. Thompson, *Phys. Rev. Lett.* 54, 1325 (1985).
6. C. E. Krohn and A. H. Thompson, *Phys. Rev. B* 33, 6366 (1986).
7. S. J. Gregg and K. S. W. Sing, "Adsorption, Surface Area and Porosity." Academic Press, London, 1982.
8. A. W. Adamson, "Physical Chemistry of Surfaces." Wiley, New York, 1975.
9. M. M. Dubinin, "Capillary Phenomena and the Information on the Pore Structure of Adsorbents," pp. 100–125. Khimia, Leningrad, 1980.
10. S. J. Sze and T. Y. Lee, *Phys. Rev. B* 51, 8709 (1995).
11. C. K. Lee, A. S. T. Chiang, and C. S. Tsay, "Porous Ceramic Materials: Fabrication, Characterization, Applications," (D.-M. Liu, Ed.) *Trans. Tech. Pub. Ltd.*, Uetikon a. S, Switzerland, 1996.
12. W. W. L. Jr., P. Schmidt-Winkel, D. Zhao, P. Feng, and G. D. Stucky, *Langmuir* 15, 5403 (1999).
13. I. Langmuir, *J. Am. Chem. Soc.* 40, 1361 (1918).
14. S. Brunauer, P. H. Emmett, and E. Teller, *J. Am. Chem. Soc.* 60, 309 (1938).
15. W. D. Machin, *J. Chem. Soc. Faraday Trans.* 88, 729 (1992).
16. K. Morishige and M. Shikimi, *J. Chem. Phys.* 108, 7821 (1998).
17. A. V. Neimark, P. I. Ravikovich, M. Grun, F. Schuth, and K. K. Unger, *J. Colloid Interface Sci.* 207, 159 (1998).
18. G. Mason, *Proc. R. Soc. London, Ser. A* 415, 453 (1988).
19. L. Sarkisov and P. A. Monson, *Langmuir* 16, 9857 (2000).
20. G. S. Heffelfinger, F. van Swol, and K. E. Gubbins, *J. Chem. Phys.* 89, 5202 (1988).
21. R. Evans and P. Tarazona, *Phys. Rev. Lett.* 52, 557 (1984).
22. M. B. Marconi and U. F. van Swol, *Europhys. Lett.* 8, 531 (1989).
23. N. A. Seaton, *Chem. Eng. Sci.* 46, 1895 (1991).
24. R. Prirard, S. Blancher, F. Brouers, and J. P. Pirard, *J. Mater. Res.* 10, 2114 (1995).
25. V. D. Borman, A. M. Grehov, and V. I. Troyan, *J. Exp. Theor. Phys.* 91, 170 (2000).
26. A. Guinier, G. Fournet, C. B. Walker, and K. L. Yudovich, "Small-Angle Scattering of X-rays." Wiley, New York, 1955.
27. L. A. Feigin and D. I. Svergun, "Structure Analysis by Small-Angle X-ray and Neutrons Scattering." Plenum, New York, 1987.
28. M. D. N. Boccara, Ed., "Physics of Finely Divided Matter." Springer-Verlag, Berlin, 1985.
29. P. Levitz and D. Tchoubar, *J. Phys. I* 2, 771 (1992).
30. P. Spanne, J. F. Thovert, C. J. Jacquin, W. B. Lindquist, K. W. Jones, and P. M. Adler, *Phys. Rev. Lett.* 73, 2001 (1994).
31. B. B. Mandelbrot, "The Fractal Geometry of Nature." Freeman, San Francisco, 1982.
32. D. Avnir, Ed., "The Fractal Approach to Heterogeneous Chemistry: Surfaces, Colloids, Polymers." Wiley, New York, 1989.

33. H. D. Bale and P. W. Schmidt, *Phys. Rev. Lett.* 53, 596 (1984).
34. P. zen Wong, J. Howard, and J.-S. Lin, *Phys. Rev. Lett.* 57, 637 (1986).
35. D. W. Schaefer and K. D. Keefer, *Phys. Rev. Lett.* 56, 2199 (1986).
36. D. W. Schaefer, B. C. Bunker, and J. P. Wilcoxon, *Phys. Rev. Lett.* 58, 284 (1987).
37. A. Hohr, H.-B. Neumann, P. W. Schmidt, P. Pfeifer, and D. Avnir, *Phys. Rev. B* 38, 1462 (1988).
38. M. Sahimi, *Rev. Mod. Phys.* 65, 1393 (1993).
39. J. C. Tarczon, A. H. Thompson, W. A. Ellingson, and W. P. Halperin, "Transport and Relaxation in Random Materials." World Scientific, Philadelphia, 1986.
40. W. P. Halperin, F. D'Orazio, S. Bhattacharja, and C. J. Tarczon, "Molecular Dynamics in Restricted Geometries." Wiley, New York, 1989.
41. F. D'Orazio, J. C. Tarczon, W. P. Halperin, K. Eguchi, and T. Mizusaki, *J. Appl. Phys.* 65, 742 (1989).
42. J. P. Butler, R. W. Mair, S. Patz, D. Hoffmann, M. I. Hrovat, R. A. Rogers, G. P. Topulos, and R. L. Walsworth, "Measuring Surface-Area-to-Volume Ratios in Soft Porous Materials Using Laser-Polarized Xenon Interphase Exchange NMR." Cond-mat/0108344.
43. F. Stallmach and J. Karger, *Adsorption* 5, 117 (1999).
44. J. Karger, H. Pfeifer, and W. Heink, *Adv. Magn. Reson.* 12, 1 (1988).
45. P. T. Callaghan, "Principles of Nuclear Magnetic Resonance Microscopy." Clarendon Press, Oxford, 1991.
46. R. Kimmich, "NMR Tomography, Diffusimetry and Relaxometry." Springer-Verlag, Berlin, 1997.
47. D. J. Bergman and K.-J. Dunn, *Phys. Rev. B* 50, 9153 (1994).
48. J. Karger and W. Heink, *J. Magn. Reson.* 51, 1 (1983).
49. P. T. Callaghan, *J. Magn. Reson. Ser. A* 113, 53 (1995).
50. P. P. Mitra and P. N. Sen, *Phys. Rev. B* 45, 143 (1992).
51. P. P. Mitra, P. N. Sen, and L. M. Schwarz, *Phys. Rev. B* 47, 8565 (1993).
52. N. K. Bar, S. Ernst, J. Karger, H. B. Schwarz, and J. Weitkamp, *Microporous Mater.* 6, 355 (1996).
53. P. N. Sen and M. D. Hurliman, *J. Chem. Phys.* 101, 5423 (1994).
54. J. Karger, N. K. Bar, W. Heink, H. Pfeifer, and G. Seiffert, *Z. Naturforsch.* 50a, 186 (1995).
55. I. M. Abdulagatov, S. N. Emirov, T. A. Tsomaeva, K. h. A. Gairbekov, and S. Y. Askerov, *Teplofizika Vysokikh Temperatur* 36, 401 (1998).
56. A. Gutina, E. Axelrod, A. Puzenko, E. Rysiakiewicz-Pasek, N. Kozlovich, and Y. Feldman, *J. Non-Cryst. Solids* 235–237, 302 (1998).
57. K. Yamamoto and H. Namikawa, *J. Ceram. Soc. Jpn.* 101, 174 (1993).
58. K. Marczuk, L. P. Prokopovich, Y. O. Roizin, E. Rysiakiewicz-Pasek, and V. N. Sviridov, *Proc. SPIE-Int. Soc. Opt. Eng.* 1513, 291 (1991).
59. G. W. Scherer, *J. Am. Ceram. Soc.* 69, 473 (1986).
60. T. H. Elmer, *Am. Ceram. Soc. Bull.* 62, 513 (1983).
61. W. Heyer, F. Wolf, and B. Kluge, *Chem. Tech.* 30, 96 (1978).
62. A. Nittke, P. Esquinazi, and A. Burin, *Phys. Rev. B* 58, 5374 (1998).
63. S. Morimoto, "Porous Ceramic Materials: Fabrication, Characterization, Applications," (D.-M. Liu, Ed.) *Trans. Tech. Pub. Ltd.*, Uetikon a. S, Switzerland, 1996.
64. X. Huang, *J. Non-Cryst. Solids* 112, 58 (1989).
65. B. I. Venzel and L. G. Svatovskaya, *Opt. Appl.* 24, 5 (1994).
66. A. M. M. Santos and W. L. Vasconcelos, *J. Non-Cryst. Solids* 273, 145 (2000).
67. P. Dieudonne, S. Calas, C. Fehr, J. Primera, T. Woignier, P. Delord, and J. Phalippou, *J. Phys. IV* 10, 73 (2000).
68. S. Morimoto, *J. Ceram. Soc. Jpn.* 98, 1291 (1990).
69. H. Hosono and Y. Abe, *J. Non-Cryst. Solids* 190, 185 (1995).
70. Y. Moriya, F. Nishikawa, Y. Kondo, N. Yamaguchi, and R. Hino, *J. Ceram. Soc. Jpn.* 103, 737 (1995).
71. H. Hosono, Y. Sakai, and Y. Abe, *J. Non-Cryst. Solids* 139, 90 (1992).
72. A. R. Boccaccini, P. Veronesi, and C. Leonelli, *J. Eur. Ceram. Soc.* 21, 1073 (2001).
73. O. V. Mazurin and E. A. Porai-Koshits, Eds., "Phase Separation in Glass." Amsterdam, 1984.
74. B. G. Varshal, Ed., "Two-Phase Glasses: Structure, Properties, Application." Nauka, Leningrad, 1991.
75. A. I. Venzel, G. P. Roskova, and T. S. Tsehomskaya, "Porous Glasses: Process of Formation, Structure, and Some Properties." Nauka, St. Petersburg, 1998.
76. G. S. Iannacchione, G. P. Crawford, S. Zumer, J. W. Doane, and D. Finotello, *Phys. Rev. Lett.* 71, 2595 (1993).
77. L. C. Klein and R. H. Woodman, "Porous Ceramic Materials: Fabrication, Characterization, Applications," (D.-M. Liu, Ed.) *Trans. Tech. Pub. Ltd.*, Uetikon a. S, Switzerland, 1996.
78. D. I. Dos-Santos, P. C. Ventura, and M. A. Aegerter, "Proceedings of NATO Advanced Study Institute," Dordrecht, the Netherlands, 1985, p. 698.
79. P. C. Santos-Ventura, D. I. Dos-Santos, and M. A. Aegerter, "Proceedings of NATO Advanced Study Institute," Dordrecht, the Netherlands, 1985, p. 697.
80. A. V. Alexeev-Popov, Y. O. Roizin, E. Rysiakiewicz-Pasek, and K. Marczuk, *Opt. Mater.* 2, 249 (1993).
81. T. Yazawa, "Porous Ceramic Materials: Fabrication, Characterization, Applications," (D.-M. Liu, Ed.) *Trans. Tech. Pub. Ltd.*, Uetikon a. S, Switzerland, 1996.
82. www.corning.com.
83. P. Levitz, G. Ehret, S. K. Sinha, and J. M. Drake, *J. Chem. Phys.* 95, 6151 (1991).
84. A. Puzenko, N. Kozlovich, A. Gutina, and Y. Feldman, *Phys. Rev. B* 60, 14348 (1999).
85. P. Wiltzius, F. S. Bates, S. B. Dierker, and G. D. Wignall, *Phys. Rev. A* 36, 2991 (1987).
86. M. Agamalian, J. M. Drake, S. K. Sinha, and J. D. Axe, *Phys. Rev. E* 55, 3021 (1997).
87. F. M. Aliev and K. S. Pozhivilko, *Sov. Fizika Tverdogo Tela* 30, 2343 (1988).
88. A. L. Dawidowicz and S. Pikus, *Appl. Surf. Sci.* 17, 45 (1983).
89. F. Katsaros, P. Makri, A. Mitropoulos, N. Kanellopoulos, U. Keiderling, and A. Wiedenmann, *Physica B* 234–236, 402 (1997).
90. H. P. Hood and M. E. Nordberg, U.S. Patent 2,106,744, 1938; U.S. Patent 2,286,275, 1942.
91. U. Even, K. Rademann, J. Jorther, N. Manor, and R. Reisfeld, *Phys. Rev. Lett.* 52, 2164 (1984).
92. N. F. Berk, *Phys. Rev. Lett.* 58, 2718 (1987).
93. J. W. Cahn, *J. Chem. Phys.* 42, 93 (1965).
94. W. Haller, *J. Chem. Phys.* 42, 686 (1965).
95. W. Haller and P. B. Macedo, *Phys. Chem. Glasses* 9, 153 (1968).
96. D. Pines and D. Huppert, *Chem. Phys. Lett.* 156, 223 (1989).
97. J. M. Drake, L. N. Yacullo, P. Levitz, and J. Klafter, *J. Phys. Chem.* 98, 380 (1994).
98. S. Henning and L. Svensson, *Phys. Sci.* 23, 697 (1981).
99. A. Fontana, M. Montagna, L. Righetti, F. Rossi, G. Cicognani, A. J. Dianoux, and F. Terki, *J. Non-Cryst. Solids* 280, 217 (2001).
100. G. Cicognani, A. J. Dianoux, A. Fontana, F. Rossi, M. Montagna, T. Scopigno, J. Pelous, F. Terki, J. N. Pilliez, and T. Woignier, *Philos. Mag. B* 79, 2091 (1999).
101. I. V. Markichev, A. Y. Muzichka, I. Natkanets, and E. F. Sheka, *Zh. Strukt. Khim.* 34, 29 (1993).
102. M. Foret, A. Chougrani, R. Vacher, and J. Pelous, *J. Phys. IV* 2, 135 (1992).
103. J. H. Root, W. J. L. Buyers, J. H. Page, D. W. Schaefer, and C. J. Brinker, "MRS Symposium," Pittsburgh, 1990, pp. 379–384.
104. R. Vacher and E. Courtens, *Proc. IEEE Ultrasonic Symp.* 2, 1237 (1989).

105. G. Reichenauer, J. Fricke, and U. Buchenau, "Low Frequency Vibrations in Silica Aerogels." Berlin, Germany, 1989.
106. H. Conrad, G. Reichenauer, and J. Fricke, "Crossover in the Vibrational Density of States of Silica Aerogels." Berlin, West Germany, 1989.
107. R. Vacher and E. Courtens, *Phys. Scr.* T29, 239 (1989).
108. E. Courtens and R. Vacher, *Proc. R. Soc. London, Ser. A* 423, 55 (1989).
109. A. Hasmy, M. Foret, E. Anglaret, J. Pelous, and R. Jullien, "Small-Angle Neutron Scattering of Aerogels: Simulations and Experiments." Cond-mat/9504055.
110. D. Staufer and R. B. Pandey, *J. Phys. A* 25, L1079 (1992).
111. R. Salazar, R. Toral, and A. Chakrabarti, *J. Sol-Gel Sci. Technol.* 15, 175 (1999).
112. P. B. Price and R. M. Walker, *Phys. Rev. Lett.* 8, 217 (1962).
113. P. B. Price and R. M. Walker, *J. Appl. Phys.* 33, 3407 (1962).
114. C. P. Bean, M. V. Doyle, and G. Entine, *J. Appl. Phys.* 41, 1454 (1970).
115. P. Apel, "Polymeric Materials Research with Cycritrons," pp. 136–143. World Scienti.c, Cape Town, South Africa, 1996.
116. Y. T. Oganessian, P. Y. Apel, A. Y. Didyk, S. N. Dmitriev, and G. G. Gulbekian, "Applied Research with Cyclotron Beams at FLNR JINR," Japan, 1996, pp. 197–200.
117. T. K. Rostoctseva, C. L. Bashford, G. M. Alder, G. H. Hill, C. McGiffert, P. Y. Apel, G. Lowe, and C. A. Pasternak, *J. Membr. Sci.* 151, 29 (1996).
118. E. Ferain and R. Legras, *Nucl. Instrum. Methods Phys. Res.* 131, 97 (1997).
119. I. Chlebny, B. Doudin, and J. P. Ansermet, *Nanostruct. Mater.* 2, 637 (1993).
120. E. Ferain and R. Legras, *Nucl. Instrum. Methods Phys. Res.* 174, 116 (2001).
121. V. N. Bogomolov, V. A. Gordeev, Y. A. Kumzerov, M. I. Martinov, and V. A. Fokin, *Sov. Tech. Phys. Lett.* 8, 599 (1982).
122. D. W. Breck, "Zeolite Molecular Sieves." Wiley, New York, 1974.
123. R. M. Barrer, "Zeolites and Clay Minerals as Sorbents and Molecular Sieves." Academic Press, London, 1978.
124. J. C. J. H. Van Bekkam and E. M. Flanigen, Eds., "Introduction to Zeolite Science and Practice." Elsevier, Amsterdam, 1991.
125. J. M. Newsam, *Science* 231, 1093 (1986).
126. T. E. Gier, X. Bu, P. Feng, and G. D. Stucky, *Nature* 395, 154 (1998).
127. P. Feng, X. Bu, and G. D. Stucky, *Nature* 388, 735 (1997).
128. X. Bu, P. Feng, and G. Stucky, *Science* 278, 2080 (1997).
129. A. Kuperman, S. Nadimi, S. Oliver, G. A. Ozin, J. M. Gerces, and M. M. Olken, *Nature* 365, 239 (1993).
130. V. N. Bogomolov, *Sov. Phys. Usp.* 21, 77 (1978).
131. W. M. Meier and D. H. Olson, "Atlas of Zeolite Structure Types," 2nd ed. Butterworths, London, 1988.
132. C. T. Kresge, M. E. Leonovicz, W. J. Roth, J. C. Vartuli, and J. S. Beck, *Nature* 359, 710 (1992).
133. A. Monnier, F. Schuth, Q. Huo, D. Kumar, D. Margolese, R. S. Maxwell, G. D. Stucky, M. Krishnamutry, P. Petroff, A. Firouzi, M. Janicke, and B. F. Chmelka, *Science* 261, 1299 (1993).
134. J. S. Beck, J. C. Vartuli, W. J. Roth, M. E. Leonowicz, C. T. Kresge, K. D. Schmitt, C. T. W. Chu, D. H. Olson, E. E. Sheppard, S. B. McCullen, J. B. Higgins, and J. L. Schlender, *J. Am. Chem. Soc.* 114, 10834 (1992).
135. M. W. Maddox, J. P. Olivier, and K. E. Gubbins, *Langmuir* 13, 1737 (1997).
136. A. Sayari, M. Kruk, M. Jaroniec, and I. L. Moudrakovski, *Adv. Mater.* 10, 1376 (1998).
137. M. Grun, I. Lauer, and K. Unger, *Adv. Mater.* 9, 254 (1997).
138. D. Zhao, J. Feng, Q. Huo, N. Melosh, G. H. Fredricson, B. F. Chmelka, and G. D. Stucky, *Science* 279, 548 (1997).
139. P. Yang, D. Zhao, D. I. Margolese, B. F. Chmelka, and G. D. Stucky, *Nature* 396, 152 (1998).
140. Y. Sakamoto, M. Kaneda, O. Terasaki, D. Zhao, J.-M. Kim, G. D. Stucky, and R. Ryoo, *Nature* 408, 449 (2000).
141. W. W. Lukens Jr. and G. D. Stucky, *Chem. Mater.* 14, 1665 (2002).
142. F. Kleitz, F. Marlow, G. D. Stucky, and F. Schuth, *Chem. Mater.* 13, 3587 (2001).
143. P. Schmidt-Winkel, W. W. Lukens Jr., P. Yang, J. S. Lettow, J. Y. Ying, and G. D. Stucky, *Chem. Mater.* 12, 686 (2000).
144. J. S. Lettow, Y. J. Han, P. Schmidt-Winkel, P. Yang, D. Zhao, G. D. Stucky, and J. Y. Ying, *Langmuir* 16, 8291 (2000).
145. H. Yang, A. Kuperman, N. Coombs, S. MamicheAfara, and G. A. Ozin, *Nature* 379, 703 (1996).
146. J. V. Sanders, *Nature* 204, 1151 (1964).
147. N. D. Deniskina, D. V. Kalinin, and L. K. Kazantseva, "Precious Opals: Synthesis and Genesis in Nature." Nauka, Novosibirsk, 1980.
148. M. Henry, *Recherche* 118, 24 (1981).
149. J. B. Jones, J. V. Sanders, and E. R. Segnit, *Nature* 204, 990 (1964).
150. V. G. Balakirev, V. N. Bogomolov, V. V. Zhuravlev, Y. A. Kumzerov, V. P. Petranovsky, S. G. Romanov, and L. A. Samoilovich, *Crystallogr. Rep.* 38, 348 (1993).
151. S. G. Romanov and C. M. Sotomayor-Torres, "Handbook of Nanostructured Materials and Nanotechnology," pp. 231–323. Academic Press, 2000.
152. Y. A. Kumzerov, "Nanostructured Films and Coatings," pp. 63–76. Kluwer Academic Publishers, Dordrecht, the Netherlands, 2000.
153. V. N. Bogomolov, Y. A. Kumzerov, S. G. Romanov, and V. V. Zhuravlev, *Physica C* 208, 371 (1993).
154. D. Elwell, "Man-made Gemstones." Ellis Horwood, Chichester, 1979.
155. R. W. J. Scott, C. M. Yang, N. Coombs, G. A. Ozin, and D. E. Williams, *Adv. Mater.* 13, 1468 (2001).
156. R. Mayoral, J. Requena, J. S. Moya, C. Lopez, A. Cintas, and H. Miguez, *Adv. Mater.* 9, 257 (1997).
157. V. N. Bogomolov, D. A. Kurdiukov, L. S. Parfenieva, A. V. Prokofiev, S. M. Samoilovich, I. A. Smirnov, A. Jezovsky, J. Mucha, and H. Miserick, *Phys. Solid State* 39, 392 (1997).
158. L. Bragg and G. F. Claringbull, "Crystal Structure of Minerals." G. Bell and Sons, London, 1965.
159. N. D. Sobolev, "Introduction to Asbestos Science." Moscow, 1971.
160. H. Schreider, "Asbestos in the Natural Environment." Elsevier, Amsterdam, 1989.
161. A. B. Bjerge, B. S. Jensen, and E. Sorensen, "Theoretical and Experimental Investigations Concerning the Decomposability of Asbestos." Riso National Laboratory, Roskilde, Denmark, 1990.
162. K. K. Zoloev and B. A. Popov, Eds., "Bazhenovo Deposit of Chrysotile Asbestos." Moscow, 1985.
163. S. R. Yuassovskii, N. B. Urusov, and N. V. Mikhailov, *Rep. Acad. Sci. USSR* 165, 1119 (1965).
164. A. S. Ognev, "Weathering influence on the Properties of Chrysotile Asbestos." Sverdlovsk, 1973.
165. D. D. Thornton, M. F. Taragin, and J. C. Eisenstein, *Phys. Lett. A* 49, 235 (1974).
166. E. J. W. Whittaker, *Acta Crystallogr.* 6, 747 (1953).
167. E. O. Shnaikovskaya, *Proc. Acad. Sci. USSR* 15, 218 (1951).
168. B. E. Warren and W. L. Bragg, *Z. Kristallogr.* 76, 201 (1930).
169. B. E. Warren and K. W. Herring, *Phys. Rev.* 59, 925 (1941).
170. E. Aruya, *Miner. Mag.* 27, 188 (1945).
171. E. J. W. Whittaker, *Acta Crystallogr.* 7, 827 (1954).
172. E. J. W. Whittaker, *Acta Crystallogr.* 8, 571 (1955).
173. E. J. W. Whittaker, *Acta Crystallogr.* 8, 261 (1955).
174. E. J. W. Whittaker, *Acta Crystallogr.* 8, 265 (1955).
175. E. J. W. Whittaker, *Acta Crystallogr.* 8, 726 (1955).
176. E. J. W. Whittaker, *Acta Crystallogr.* 9, 85 (1956).
177. E. J. W. Whittaker, *Acta Crystallogr.* 9, 862 (1956).
178. E. J. W. Whittaker, *Acta Crystallogr.* 9, 865 (1956).
179. E. J. W. Whittaker, 10, 149 (1957).
180. K. Yada, *Acta Crystallogr.* 23, 704 (1967).

181. N. N. Padurow, *Acta Crystallogr.* 3, 200 (1950).
182. G. W. Brindley, "X-ray Identification of Crystal Structures of Clay Minerals." Mineralogical Society, London, 1952.
183. H. Jagodzinski and G. Kunze, *Neues Jahrb. Mineral.* 95, 137 (1954).
184. J. E. W. Whittaker and J. Zussmann, *Miner. Mag.* 31, 107 (1956).
185. F. L. Pundsak, *J. Phys. Chem.* 60, 361 (1956).
186. F. L. Pundsak, *J. Phys. Chem.* 65, 30 (1961).
187. A. W. Naumann and W. H. Dresher, *Am. Mineral.* 51, 711 (1966).
188. K. Yada, *Acta Crystallogr. Sect. A* 27, 659 (1971).
189. A. E. Charola, S. Z. Lewin, and M. J. Yoder, *J. Cryst. Growth* 39, 571 (1977).
190. S. Iijima, *Nature* 354, 56 (1991).
191. M. S. Dresselhaus, G. Dresselhaus, and P. G. Eklund, "Science of Fullerenes and Carbon Nanotubes." Academic, New York, 1996.
192. T. W. Ebbesen, "Carbon Nanotubes: Preparation and Properties." CRC, Boca Raton, FL, 1997.
193. R. Saito, G. Dresselhaus, and M. S. Dresselhaus, "Physical Properties of Carbon Nanotubes." Imperial College Press, London, 1998.
194. M. S. Dresselhaus, G. Dresselhaus, and P. Avouris, Eds., "Carbon Nanotubes: Synthesis, Structure, Properties and Applications." Springer-Verlag, Berlin, 2000.
195. K. F. K. Tanaka and T. Yamabe, Ed., "The Science and Technology of Carbon Nanotubes." Elsevier, Amsterdam, 1999.
196. H. S. Nalwa, Ed., "Handbook of Nanostructured Materials and Nanotechnology." Academic Press, San Diego, 2000.
197. T. W. Ebbesen, *Annu. Rev. Mater.* 24, 235 (1994).
198. P. M. Ajayan and T. W. Ebbesen, *Rep. Progr. Phys.* 60, 1025 (1997).
199. P. M. Ajayan and O. Z. Zhou, "Application of Carbon Nanotubes," pp. 391–425. Springer-Verlag, Berlin/Heidelberg, 2001.
200. N. Wang, Z. K. Tang, G. D. Li, and J. S. Chen, *Nature* 408, 50 (2000).
201. P. M. Ajayan and S. Iijima, *Nature* 361, 333 (1993).
202. S. C. Tsang, Y. K. Chen, P. J. F. Harris, and M. L. H. Green, *Nature* 372, 159 (1994).
203. E. Dujardin, T. W. Ebbesen, T. Hiura, and K. Tanigaki, *Science* 265, 1850 (1994).
204. D. Ugarte, T. Stockly, J. M. Bonard, A. Chatelain, and W. A. D. Heer, "Capillarity in Carbon Nanotubes," p. 128. Elsevier, 1999.
205. J. Sloan and M. L. H. Green, "Handbook of Nanostructured Materials and Nanotechnology," Vol. 5, p. 407. Academic Press, San Diego 2000.
206. R. Tenne, L. Margulis, M. Genut, and G. Hodes, *Nature* 360, 444 (1992).
207. A. Rubio, J. L. Gorkill, and M. L. Cohen, *Phys. Rev. B* 49, 5081 (1994).
208. X. Blaise, A. Rubio, S. G. Louie, and M. L. Cohen, *Eur. Lett.* 28, 335 (1994).
209. Y. Miamoto, A. Rubio, and S. G. Louie, *Phys. Rev. B* 50, 18360 (1994).
210. Y. R. Hacohen, E. Grunbaum, R. Tenne, J. Sloan, and J. L. Hutchison, *Nature* 395, 336 (1998).
211. A. L. Ivanovskii, *Usp. Khim.* 71, 203 (2002).
212. P. G. D. Gennes, *Rev. Mod. Phys.* 57, 827 (1985).
213. J. S. Rowlinson and B. Wodom, "Molecular Theory of Capillarity." Clarendon Press, Oxford, 1982.
214. E. Rysiakiewicz-Pasek, *J. Electrostatics* 51–52, 173 (2001).
215. J. C. M. Li, *J. Alloys Compd.* 310, 24 (2000).
216. Y. Feldman, Y. Ryabov, A. Puzenko, A. Gutina, and N. Kozlovich, "American Institute of Physics Conference Proceedings," 1999, pp. 236–242.
217. J. M. Zanotti, M. C. Bellissent-Funel, and S. H. Chen, *Phys. Rev. E: Stat. Phys.* 59, 3084 (1999).
218. M. C. Bellissent-Funel, J. Lal, and L. Bosio, *J. Chem. Phys.* 98, 4246 (1993).
219. H. Pfeifer, W. Oehme, and H. Siegel, *Annal. Phys.* 42, 496 (1985).
220. S. Mitra, R. Mukhopadhyay, I. Tsukushi, and S. Ikeda, *J. Phys.: Condens. Matter.* 13, 8455 (2001).
221. J. Dore, *Chem. Phys.* 258, 327 (2000).
222. C. M. B. Line and G. J. Kearley, *J. Chem. Phys.* 112, 9058 (2000).
223. D. Wallacher, R. Ackermann, P. Huber, M. Enderle, and K. Knorr, *Phys. Rev. B*: *Condens. Matter* 64, 184203 (2001).
224. D. Wallacher, P. Huber, and K. Knorr, *J. Low Temp. Phys.* 122, 313 (2001).
225. H. Wendt and R. Richert, *Symp. Mater. Res. Soc.* p. 15, Warrendale, PA, 1999.
226. P. Huber and K. Knorr, *Phys. Rev. B: Condens. Matter* 60, 12657 (1999).
227. P. Huber, D. Wallacher, and K. Knorr, *Phys. Rev. B* 60, 12666 (1999).
228. E. Molz, A. P. Y. Wong, M. H. W. Chan, and J. R. Beamish, *J. Am. Chem. Soc.* 46, 596 (1924).
229. P. E. Sokol, R. T. Azuah, M. R. Gibbs, and S. M. Bennington, *J. Low Temp. Phys.* 103, 23 (1996).
230. V. E. Syvokon, K. Ooyama, M. Fukuda, T. Obata, T. Igarashi, and M. Kubota, *Physica B* 284–288, 113 (2000).
231. M. H. W. Chan, *Acad. Sci. Czech. Republic* 46, 2915 (1996).
232. N. Mulders, J. Mas, S. Kim, J. Yoon, and M. H. W. Chan, *J. Phys.: Condens. Matter* 8, 9609 (1996).
233. D. N. Bittner and E. D. Adams, *J. Low Temp. Phys.* 97, 519 (1994).
234. M. H. W. Chan, *Physica B* 169, 135 (1991).
235. G. G. Ihas and G. F. Spencer, *Physica B* 165–166, 567 (1990).
236. J. R. Beamish and N. Mulders, "American Institute of Physics Conference Proceedings," 1989, pp. 182–190.
237. M. H. W. Chan, "American Institute of Physics Conference Proceedings," 1989, pp. 170–178.
238. M. Shimoda, T. Mizusaki, T. Suzuki, A. Hirai, and K. Eguchi, "Proceedings of the 17th International Conference on Low Temperature Physics," Amsterdam, 1984, Vol. 2, pp. 1223–1224.
239. M. Shimoda, T. Mizusaki, T. Suzuki, A. Hirai, and K. Eguchi, *Phys. Lett. A* 102A, 426 (1984).
240. W. S. Truscott and D. F. Brewer, *J. Phys.* (*Paris*) 41, 267 (1980).
241. R. M. Dimeo, P. E. Sokol, C. R. Anderson, W. G. Stirling, and M. A. Adams, *J. Low Temp. Phys.* 113, 369 (1998).
242. C. W. Kiewiet, H. E. Hall, and J. D. Reppy, *Phys. Rev. Lett.* 35, 1286 (1975).
243. M. Vilfan, T. Apih, A. Gregorovic, B. Zalar, G. Lahajnar, S. Zumer, G. Hinze, R. Bohmer, and G. Althoff, *Magn. Reson. Imaging* 19, 433 (2001).
244. F. Grinberg and R. Kimmich, *Magn. Reson. Imaging* 19, 401 (2001).
245. J. Werner, K. Otto, D. Enke, G. Pelzl, F. Janowski, and H. Kresse, *Liq. Cryst.* 27, 1295 (2000).
246. S. Kralj, A. Zidansek, G. lahajnar, S. Zumer, and R. Blinc, *Phys. Rev. E* 62, 718 (2000).
247. G. P. Sinha, B. Batalla, and F. M. Aliev, *Symp. Mater. Res. Soc.* p. 151, Warrendale, PA, 1998.
248. Y. P. Panarin, C. Rosenblatt, and F. M. Aliev, *Phys. Rev. Lett.* 81, 2699 (1998).
249. S. Kralj, A. Zidansek, G. Lahajnar, S. Zumer, and R. Blinc, *Phys. Rev. E* 57, 3021 (1998).
250. F. M. Aliev and I. V. Plechakov, "MRS Symposium," Pittsburgh, 1997, pp. 177–182.
251. F. M. Aliev and G. P. Sinha, "MRS Symposium," Pittsburgh, 1996, pp. 413–418.
252. F. M. Aliev and V. V. Nadtotchi, "MRS Symposium," Pittsburgh, 1996, pp. 125–130.
253. F. M. Aliev and J. Kelly, *Ferroelectrics* 151, 263 (1994).
254. T. Bezrodnaya, V. Mel'nik, and K. Nelipovich, *J. Mol. Struct.* 596, 55 (2001).
255. T. Bezrodnaya, T. Gavrilko, I. Gnatyuk, O. Yaroshchuk, G. Puchkovskaya, J. Baran, H. Ratajczak, and J. I. Kukielski, *Tech. Univ. Wrocław. Optica Applicata* 30, 585 (2000).
256. M. Sliwinska-Bartkowiak, G. Dudziak, R. Sikorski, R. Gras, R. Radhakrishnan, and K. E. Gubbins, *J. Chem. Phys.* 114, 950 (2001).

257. D. Majolino, V. Crupi, G. Maisano, P. Migliardo, and V. Venuti, "American Institute of Physics Conference Proceedings," 2000, pp. 94–97.
258. V. Crupi, D. Majolino, G. Maisano, P. Migliardo, and V. Venuti, *Philos. Mag. B* 79, 1871 (1999).
259. P. N. Prasad, N. D. Kumar, L. Lal, and M. P. Joshi, "Electrical," pp. 227–235. Warrendale, PA, 1998.
260. V. Crupi, D. Majolino, P. Migliardo, and V. Venuti, *J. Mol. Struct.* 448, 255 (1998).
261. D. e. A. J. Gorgatti, M. A. Paoli, and O. L. Alves, *Synth. Met.* 84, 107 (1997).
262. X. Yan, C. Streck, and R. Richert, "MRS Symposium," Pittsburgh, 1997, pp. 33–44.
263. E. V. Colla, E. Y. Koroleva, Y. A. Kumzerov, B. N. Savenko, and S. B. Vakhrushev, *Ferroelectrics Lett.* 8, L203 (1996).
264. E. V. Colla, E. Y. Koroleva, Y. A. Kumzerov, B. N. Savenko, and S. B. Vakhrushev, *Ferroelectrics Lett.* 20, 143 (1996).
265. E. V. Colla, A. V. Fokin, and Yu. A. Kumzerov, *Solid State Commun.* 103, 127 (1997).
266. M. P. Joshi, L. M. Kukreja, and K. C. Rustagi, *Appl. Phys. A* 65, 5 (1997).
267. D. A. Kurdiukov, Y. A. Kumzerov, A. V. Medvedev, A. B. Pevtsov, V. N. Semkin, and M. V. Zamoryanskaya, "Abstracts of the 5th Biennial International Workshop on Fullerenes and Atomic Clusters," 2001.
268. N. V. Drichko, D. A. Kurdiukov, Y. A. Kumzerov, B. V. Petrov, V. N. Semkin, R. Almairac, P. Diedonne, Y. Errammach, and F. Rachdi, "Abstracts of the 5th Biennial International Workshop on Fullerenes and Atomic Clusters," 2001.
269. P. M. Ajayan, O. Stephan, P. Redlich, and C. Colliex, *Nature* 375, 564 (1995).
270. V. N. Bogomolov, V. K. Krivosheev, R. S. h. Malkovich, and F. A. Chudnovskii, *Fiz. Tverdogo Tela* 11, 3648 (1969).
271. V. N. Bogomolov, R. S. h. Malkovich, and F. A. Chudnovskii, *Fiz. Tverdogo Tela* 11, 2835 (1969).
272. V. N. Bogomolov, N. A. Klushin, N. M. Okuneva, E. L. Plachenova, V. I. Pogrebnoi, and F. A. Chudnovskii, *Fiz. Tverdogo Tela* 13, 1499 (1971).
273. D. Michel, B. F. Borisov, E. V. Charnaya, W. D. Hoffmann, P. G. Plotnikov, and Y. A. Kumzerov, *Nanostruct. Mater.* 12, 515 (1999).
274. B. F. Borisov, E. V. Charnaya, P. G. Plotnikov, W.-D. Hoffman, D. Michel, Y. A. Kumzerov, C. Tien, and C.-S. Wur, *Phys. Rev. B* 58, 5329 (1998).
275. E. V. Charnaya, C. Tien, K. J. Lin, and Y. A. Kumzerov, *Phys. Rev. B* 58, 11089 (1998).
276. Y. A. Kumzerov, A. A. Naberezhnov, S. B. Vakhrushev, and B. N. Savenko, *Phys. Rev. B* 52, 4772 (1995).
277. K. M. Unruh, J. F. Sheehan, T. E. Huber, and C. A. Huber, *Nanostruct. Mater.* 3, 425 (1993).
278. K. M. Unruh, T. E. Huber, and C. A. Huber, *Phys. Rev. B: Condens. Matter* 48, 9021 (1993).
279. W. J. Standish, "Proceedings of the 9th AIRAPT International High Pressure Conference," New York, 1984, p. 71.
280. W. J. Standish and R. L. Pompi, *Phys. Rev. B: Condens. Matter* 21, 5185 (1980).
281. N. K. Hindley and J. H. P. Watson, *Phys. Rev. A* 183, 525 (1969).
282. J. H. P. Watson and R. M. Hawk, *Solid State Commun.* 9, 1993 (1971).
283. J. H. P. Watson, "Abstracts of the 12th International Conference on Low Temperatures Physics," 1970, p. 230.
284. J. H. P. Watson, *Appl. Phys. Lett.* 16, 428 (1970).
285. V. N. Bogomolov, *Sov. Phys. Solid State* 13, 672 (1971).
286. V. N. Bogomolov and A. I. Zhadorozhnii, *Sov. Phys. Solid State* 17, 1078 (1975).
287. V. N. Bogomolov, T. I. Volkonskaya, A. I. Zadorozhnii, A. A. Kapanadze, and E. L. Lutsenko, *Sov. Phys. Solid State* 17, 1110 (1975).
288. V. N. Bogomolov and T. M. Pavlova, *Semiconductors* 29, 428 (1995).
289. Bogomolov, V. Zhuravlev, A. Zhadorozhnii, E. Kolla, and Y. Kumzerov, *JETP Lett.* 36, 365 (1982).
290. V. N. Bogomolov, L. K. Kazantseva, E. V. Kolla, and Y. A. Kumzerov, *Sov. Phys. Solid State* 29, 359 (1987).
291. Y. Kumzerov, V. Bogomolov, E. Colla, and S. Romanov, *Phys. Low-Dim. Struct.* 11–12, 129 (1994).
292. K. K. Babamuratov, V. V. Zhuravlev, Y. A. Kumzerov, S. G. Romanov, and S. A. Khachaturov, *Sov. Phys. Solid State* 35, 795 (1993).
293. V. N. Bogomolov, V. K. Krivosheev, and Y. A. Kumzerov, *Sov. Phys. Solid State* 13, 3148 (1972).
294. V. N. Bogomolov and Y. A. Kumzerov, *JETP Lett.* 21, 198 (1975).
295. V. N. Bogomolov, B. E. Kvyatkovskii, E. V. Kolla, S. A. Ktitorov, Y. A. Kumzerov, and N. M. Okuneva, *Sov. Phys. Solid State* 23, 1271 (1981).
296. V. Bogomolov, Y. Kumzerov, and V. Pimenov, *Phys. Lett. A* 86, 183 (1981).
297. V. N. Bogomolov, E. V. Kolla, Y. A. Kumzerov, N. M. Okuneva, and V. N. Prigodin, *Solid State Commun.* 35, 363 (1980).
298. V. N. Bogomolov, N. A. Klushin, and Y. A. Kumzerov, *JETP Lett.* 26, 72 (1977).
299. V. N. Bogomolov, E. V. Kolla, and Y. A. Kumzerov, *Solid State Commun.* 46, 159 (1983).
300. M. S. Ivanova, Y. A. Kumzerov, V. V. Poborchii, Y. V. Ulashkevich, and V. V. Zhuravlev, *Microporous Mater.* 4, 319 (1995).
301. M. V. Vedernikov, O. N. Urypin, B. M. Goltsman, Y. V. Ivanov, and Y. A. Kumzerov, "Proceedings of the 2001 MRS Meeting," Symposium g, 2001.
302. S. V. Zaitsev-Zotov, Y. A. Kumzerov, Y. A. Firsov, and P. Monceau, *J. Phys.: Condens. Matter* 12, L303 (2000).
303. S. V. Zaitsev-Zotov, Y. A. Kumzerov, Y. A. Firsov, and P. Monceau, "Proceedings of the 10th International Symposium on Nanostructures: Physics and Technology," 2002, p. 512.
304. V. N. Bogomolov, *Phys. Solid State* 35, 469 (1993).
305. V. N. Bogomolov, *Zeolite News Lett.* 10, 148 (1993).
306. V. N. Bogomolov, *Phys. Rev. B* 51, 17040 (1995).
307. V. N. Bogomolov, *Technical Phys.* 38, 224 (1993).
308. V. N. Bogomolov, *Sov. Phys. Tech. Phys.* 37, 79 (1992).
309. V. N. Bogomolov, *Surface: Phys. Chem. Mechanics* 9, 136 (1992).
310. A. D. Berry, R. J. Tonucci, and M. Fatemi, *Appl. Phys. Lett.* 69, 2846 (1996).
311. K. S. Pater, *Opt. Appl.* 25, 189 (1995).
312. D. G. Hendershot, D. K. Gaskill, B. L. Justus, M. Fatemi, and A. D. Berry, *Appl. Phys. Lett.* 63, 3324 (1993).
313. S. G. Romanov, V. Y. Butko, Y. A. Kumzerov, N. M. Yats, M. I. Pemble, J. R. Agger, M. W. Anderson, and C. M. Sotomayor-Torres, *Phys. Solid State* 39, 641 (1997).
314. J. C. Luong, *Superlattices Microstruct.* 4, 385 (1988).
315. N. Murase and T. Yazawa, *J. Am. Ceram. Soc.* 84, 2269 (2001).
316. N. F. Borrelli and J. C. Luong, *Proc. SPIE-Int. Soc. Opt. Eng.* 866, 104 (1988).
317. I. V. Golosovsky, I. Mirebeau, G. Andre, D. A. Kurdyukov, Y. A. Kumzerov, and S. B. Vakhrushev, *Phys. Rev. Lett.* 86, 5783 (2001).
318. V. N. Bogomolov, V. G. Golubev, N. F. Katrenko, D. A. K. A. B. Pevtsov, A. V. Prokofiev, V. V. Ratnikov, N. A. Feoktistov, and N. V. Sharenkova, *Techn. Phys. Lett.* 24, 326 (1998).
319. V. N. Bogomolov, N. A. Feoktistov, V. G. Golubev, J. L. H. D. A. K. A. B. Pevtsov, J. Sloan, and L. M. Sorokin, "Proceedings of the 10th International Symposium on Nanostructures: Physics and Technology," 1999, p. 52.
320. V. N. Bogomolov, N. A. Feoktistov, V. G. Golubev, J. L. Hutchinson, D. A. Kurdiukov, A. B. Pevtsov, R. Schwarz, J. Sloan, and L. M. Sorokin, *J. Non-Cryst. Solids* 266–269, 1021 (2000).

321. C. Diaz-Guerra, J. Piqueras, V. G. Golubev, D. A. Kurdiukov, A. B. Pevtsov, and M. V. Zamoryanskaya, *Appl. Phys. Lett.* 77, 3194 (2000).
322. C. Diaz-Guerra, D. A. Kurdiukov, J. Piqueras, V. I. Sokolov, and M. V. Zamoryanskaya, *J. Appl. Phys.* 89, 2720 (2001).
323. V. Y. Davidov, V. G. Glubev, N. F. Kartenko, D. A. Kurdiukov, A. B. Pevtsov, N. V. Sharenkova, P. Brogueira, and R. Schwarz, *Nanotechnology* 11, 291 (2000).
324. G. Gajiev, V. G. Golubev, D. A. Kurdiukov, A. B. Pevtsov, A. V. Selkin, and V. V. Travnikov, *Phys. Status Solid: B* 231, R7 (2002).
325. V. Y. Davidov, R. E. Dunin-Borkovski, V. G. Golubev, J. L. Hutchinson, N. F. Kartenko, D. A. Kurdiukov, A. B. Petsov, N. V. Sharenkova, J. Sloan, and L. M. Sorokin, *Semicond. Sci. Technol.* 16, L5 (2001).
326. V. G. Golubev, D. A. Kurdiukov, A. V. Medvedev, A. B. Pevtsov, L. M. Sorokin, and J. L. Hutchinson, *Semiconductors* 35, 1320 (2001).
327. V. G. Golubev, V. Y. Davidov, N. F. Kartenko, D. A. Kurdiukov, A. V. Medvedev, A. B. Pevtsov, A. V. Scherbakov, and E. B. Shadrin, *Appl. Phys. Lett.* 79, 2127 (2001).
328. G. D. Stucky and J. E. M. Dougall, *Science* 247, 669 (1990).
329. A. Monnier, V. Strdanov, G. Stucky, and H. Metiu, *J. Chem. Phys.* 100, 6944 (1994).
330. M. T. Janicke, C. C. Landry, S. C. Christiansen, D. Kummer, G. D. Stucky, and B. F. Chmelka, *J. Am. Chem. Soc.* 120, 6940 (1998).
331. M. J. M. Laghlan, N. Coombs, and G. A. Ozin, *Nature* 397, 681 (1999).
332. R. Leon, D. Margolese, G. Stucky, and P. M. Petroff, *Phys. Rev. B* 52, R2285 (1995).
333. Y. J. Han, J. M. Kim, and G. D. Stucky, *Chem. Mater.* 12, 2068 (2000).
334. C. S. Yang, D. D. Awshalom, and G. D. Stucky, *Chem. Mater.* 14, 1277 (2002).
335. P. Yang, A. H. Rizvi, B. Messer, B. F. Chmelka, G. M. Whitesides, and G. D. Stucky, *Adv. Mater.* 13, 427 (2001).
336. S. D. Cox and G. D. Stucky, *J. Phys. Chem.* 95, 710 (1991).
337. G. A. Ozin, I. Manners, and M. J. M. Lachlan, *Adv. Mater.* 12, 675 (2000).
338. G. A. Ozin, C.-Y. Ishii, T. Asefa, and M. J. M. Lachlan, *J. Mater. Chem.* 10, 1751 (2000).
339. T. Asefa, M. J. MacLaghlan, N. Coombs, and G. A. Ozin, *Nature* 402, 867 (1999).
340. R. M. Lago, S. C. Tsang, K. L. Lu, Y. K. Chen, and M. L. H. Green, *J. Chem. Soc., Chem. Commun.* 1355 (1995).
341. A. Chu, J. Cook, R. Heesom, J. L. Hutchinson, M. L. H. Green, and J. Sloan, *Chem. Mater.* 8, 2751 (1996).
342. Y. K. Chen, A. Chu, J. Cook, M. L. H. Green, P. J. F. Harris, R. Heesom, M. Humphries, J. Sloan, S. C. Tsang, and J. C. F. Turner, *J. Mater. Chem.* 7, 545 (1997).
343. S. Pruneanu, G. Mihailescu, and E. Indrea, "Proceedings of the 2000 International Semiconductor Conference," Piscataway, NJ, 2000, Vol. 2, p. 475.
344. Y. Peng, D. H. Qin, R. J. Zhou, and H. L. Li, *Mater. Sci. Eng., B* 77, 246 (2000).
345. X. F. Wang, J. Zhang, H. Z. Shi, Y. W. Wang, G. W. Meng, X. S. Peng, L. D. Zhang, and J. Fang, *J. Appl. Phys.* 89, 3847 (2001).
346. G. Tourillon, L. Pontonnier, J. P. Levy, and V. Langlais, *Electrochem. Solid-State Lett.* 3, 20 (2000).
347. S. A. Gusev, N. A. Korotkova, D. B. Rozenstein, and A. A. Fraerman, *J. Appl. Phys.* 76, 6671 (1994).
348. V. I. Klimov, "Handbook of Nanostructured Materials and Nanotechnology," Vol. 4, pp. 451–527. Academic Press, San Diego, 2000.
349. J. C. Luong and N. F. Borelli, *Mater. Res. Soc. Symp. Proc.* 144, 695 (1989).
350. V. I. Strdanov, I. Alxneip, G. D. Stucky, C. M. Reaves, and S. P. DenBaars, *J. Phys. Chem.* 102, 3341 (1998).
351. V. F. Agekyan and A. Y. Serov, *Fiz. Tverdogo Tela* 38, 122 (1996).
352. V. V. Poborchii, M. S. Ivanova, and I. A. Salamatina, *Superlattices Microstruct.* 16, 133 (1994).
353. V. V. Poborchii, *Jpn. J. Appl. Phys. Suppl.* 34, 271 (1994).
354. E. A. Muljarov, E. A. Zhukov, V. S. Dneprovskii, and Y. Masumoto, *Phys. Rev. B* 62, 7420 (2000).
355. V. V. Poborchii, V. I. Alperovich, Y. Nozue, N. Ohnishi, A. Kasuya, and O. Terasaki, *J. Phys.: Condens. Matter* 9, 5687 (1997).
356. J. C. Maxwell-Garnett, *Philos. Trans. R. Soc. London* 203, 385 (1904).
357. J. D. Jackson, "Classical Electrodynamics," Chap. 4. Wiley, New York, 1962.
358. S.-I. Lee, T. W. Noh, J. R. Gaines, K. O. Ying-Hsiang, and E. R. Kreidler, *Phys. Rev. B: Condens. Matter* 37, 2918 (1988).
359. C. A. Foss, G. L. Hornyak, J. A. Stockert, and C. R. Martin, *J. Phys. Chem.* 98, 2963 (1994).
360. R. W. Cohen, G. D. Cody, M. D. Coutts, and B. Abeles, *Phys. Rev. B* 8, 3689 (1973).
361. L. Ren-Sheng and J. Jonas, *J. Raman Spectrosc.* 32, 975 (2001).
362. V. Crupi, G. Maisano, D. Majolino, P. Migliardo, and V. Venuti, *J. Chem. Phys.* 109, 7394 (1998).
363. G. Carini, V. Crupi, G. D'Angelo, D. Majolino, P. Migliardo, and Y. B. Mel'nichenko, *J. Chem. Phys.* 107, 2292 (1997).
364. T. E. Huber and C. A. Huber, *Appl. Phys. A* 51, 137 (1990).
365. M. J. D. Low and P. L. Bartner, *Can. J. Chem.* 48, 7 (1970).
366. M. I. Abaev, V. N. Bogomolov, V. V. Bryksin, and N. A. Klushin, *Fiz. Tverdogo Tela* 13, 1578 (1971).
367. I. Gnatyuk, G. Puchkovskaya, O. Yaroshchuk, Y. Goltsov, L. Matkovskaya, J. Baran, T. Morawska-Kowal, and H. Ratajczak, *J. Mol. Struct.* 511–512, 189 (1999).
368. G. Schwalb and F. W. Deeg, *Phys. Rev. Lett.* 74, 1383 (1995).
369. S. G. Romanov, C. M. S. Torres, H. M. Yates, M. E. Pemble, V. Butko, and V. Tretijakov, *J. Appl. Phys.* 82, 380 (1997).
370. D. P. Savin, S. A. Gevelyuk, Y. A. Roizin, E. Mugenski, and I. Sokolska, *Appl. Phys. Lett.* 72, 3005 (1998).
371. M. P. Joshi and A. Ingale, *Appl. Phys. B* 72, 941 (2001).
372. V. I. Zemskii, A. V. Veresov, and A. Y. Ershov, *Opt. Spektrosk.* 81, 251 (1996).
373. V. I. Zemskii and A. V. Veresov, *Opt. Spektrosk.* 81, 796 (1996).
374. S. A. Bagnich, V. N. Bogomolov, Y. A. Kumzerov, and P. P. Perushkevich, *Phys. Solid State* 37, 1115 (1995).
375. S. A. Bagnich, *Phys. Rev. B* 60, 11374 (1999).
376. S. A. Bagnich, *Opt. Spectrosc.* 90, 375 (2001).
377. H. M. Yates, W. R. Flavell, M. E. Pemble, N. P. Johnson, S. G. Romanov, and C. M. SotomayorTorres, *J. Cryst. Growth* 170, 611 (1997).
378. M. W. Anderson, G. K. Logothetis, A, G. Taylor, N. Wallace, and H. M. Yates, *Adv. Mater. Opt. Electron.* 2, 313 (1993).
379. V. N. Bogomolov, E. L. Lutsenko, V. P. Petranovskii, and S. V. Kholodkevich, *JETP Lett.* 23, 482 (1976).
380. V. N. Bogomolov, A. I. Zadorozhnii, V. P. Petranovskii, A. V. Fokin, and S. V. Kholodkevich, *JETP Lett.* 29, 373 (1979).
381. V. N. Bogomolov, V. V. Poborchii, S. V. Kholodkevich, and S. I. Shagin, *JETP Lett.* 38, 532 (1983).
382. V. N. Bogomolov, V. P. Petranovskii, V. V. Poborchii, and S. V. Kholodkevich, *Sov. Phys. Solid State* 25, 1415 (1983).
383. V. N. Bogomolov, V. V. Poborchii, and S. V. Kholodkevich, *JETP Lett.* 42, 517 (1985).
384. V. N. Bogomolov, V. V. Poborchii, S. G. Romanov, and S. I. Shagin, *J. Phys. C* 18, L313 (1985).
385. Y. Nozue, T. Kodaira, O. Terasaki, K. Yamazaki, T. Goto, D. Watanabe, and J. M. Thomas, *J. Phys. Condens. Matter* 2, 5209 (1990).
386. V. V. Poborchii, M. S. Ivanova, V. P. Petranovskii, Y. A. Barnakov, A. Kasuya, and Y. Nishina, *Mater. Sci. Eng., A* 217/218, 129 (1996).
387. Z. Lin, Z. Wang, W. Chen, L. Lin, G. Li, Z. Liu, and H. Han, *Solid State Commun.* 100, 841 (1996).

388. V. V. Poborchii, *Solid State Commun.* 107, 513 (1998).
389. V. V. Poborchii, *J. Chem. Phys.* 114, 2707 (2001).
390. V. V. Poborchii, *Chem. Phys. Lett.* 251, 230 (1996).
391. V. N. Bogomolov, A. N. Efimov, M. S. Ivanova, V. V. Poborchii, S. G. Romanov, Y. I. Smolin, and Y. F. Shepelev, *Sov. Phys. Solid State* 34, 916 (1992).
392. V. V. Poborchii, M. Sato, and A. V. Schukarev, *Solid State Commun.* 103, 649 (1997).
393. V. V. Poborchii, G.-G. Lindner, and M. Sato, *J. Chem. Phys.* 116, 2609 (2002).
394. V. V. Poborchii, A. V. Kolobov, J. Caro, V. V. Zhuravlev, and K. Tanaka, *Chem. Phys. Lett.* 280, 17 (1997).
395. V. V. Poborchii, A. V. Kolobov, H. Oyanagi, J. Caro, V. V. Zhuravlev, and K. Tanaka, "MRS Proceedings of the 12th International Zeolite Conference," 1999.
396. V. V. Poborchii, A. V. Kolobov, J. Caro, V. V. Zhuravlev, and K. Tanaka, *Phys. Rev. Lett.* 82, 1955 (1999).
397. S. Schmitt-Rink, D. A. B. Miller, and D. S. Chemla, *Phys. Rev. B* 35, 8113 (1987).
398. L. Banyai, Y. Z. Hu, M. Lindberg, and S. W. Koch, *Phys. Rev. B* 38, 8142 (1988).
399. E. Hanamura, *Phys. Rev. B* 37, 1273 (1988).
400. B. L. Justus, R. J. Tonucci, and A. D. Berry, *Appl. Phys. Lett.* 61, 3151 (1992).
401. M. D. Dvorak, B. L. Justus, and A. D. Berry, *Opt. Commun.* 116, 149 (1995).
402. M. D. Dvorak, B. L. Justus, D. K. Gaskill, and D. G. Hendershot, *Appl. Phys. Lett.* 66, 804 (1995).
403. M. Seik-Bahae, A. A. Said, T. H. Wei, D. J. Hagan, and E. W. V. Stryland, *IEEE J. Quantum Electron.* 26, 760 (1990).
404. V. S. Dneprovskii, E. A. Zhukov, N. Y. Markova, E. A. Mulyarov, K. A. Chernoutsan, and O. A. Shalygina, *Phys. Solid State* 42, 544 (2000).
405. E. A. Zhukov, Y. Masumoto, E. A. Muljarov, and S. G. Romanov, *Solid State Commun.* 112, 575 (1999).
406. V. S. Dneprovskii, E. A. Zhukov, E. A. Muljarov, and S. G. Tikhodeev, *J. Exp. Theor. Phys.* 87, 382 (1998).
407. V. Dneprovskii, E. Zhukov, V. Karavanskii, V. Poborchii, and I. Salamatina, *Superlattices Microstruct.* 23, 1217 (1998).
408. V. Dneprovskii and E. Zhukov, *Phys. Status Solid: B* 206, 469 (1998).
409. V. Dneprovskii, V. Karavanskii, V. Poborchii, I. Salamatina, and E. Zhukov, "Compound Semiconductors," Institute of Physics Conference Series, p. 825, IOP Publishing, Bristol, 1997.
410. V. Dneprovskii, *Laser Phys.* 6, 997 (1996).
411. E. A. Mulyarov and S. G. Tikhodeev, *J. Exp. Theor. Phys.* 84, 151 (1997).
412. E. Yablonovich, *Phys. Rev. Lett.* 58, 2059 (1987).
413. S. John, *Phys. Rev. Lett.* 58, 2486 (1987).
414. V. N. Bogomolov, S. V. Gaponenko, I. N. Germanenko, A. M. Kapitonov, E. P. Petrov, N. V. Gaponenko, A. V. Prokofiev, A. N. Ponyavina, N. I. Silvanovich, and S. M. Samoilovich, *Phys. Rev. E* 55, 7619 (1997).
415. Y. A. Vlasov, V. N. Astratov, O. Z. Karimov, A. A. Kaplyanskii, V. N. Bogomolov, and A. V. Prokofiev, *Phys. Rev. B* 55, R13357 (1997).
416. Y. A. Vlasov, M. A. Kaliteevski, and V. V. Nikolaev, *Phys. Rev. B* 60, 1555 (1999).
417. Y. A. Vlasov, V. N. Astratov, A. V. Barishev, A. A. Kaplyanskii, O. Z. Karimov, and M. F. Limonov, *Phys. Rev.* 61, 5784 (2000).
418. S. G. Romanov, T. Maka, C. M. Sotomayor-Torres, M. Muller, R. Zentel, D. Cassagne, J. Manzanares-Martines, and C. Jouanin, *Phys. Rev. E* 63, 056603 (2001).
419. K. Bush and S. John, *Phys. Rev. E* 58, 3896 (1998).
420. S. Tsunekawa, Y. A. Barnakov, V. V. Poborchii, S. M. Samoilovich, A. Kasuya, and Y. Nishina, *Microporous Mater.* 8, 275 (1997).
421. D. Kang, J. E. Maclennan, N. A. Clark, A. A. Zakhidov, and R. H. Baughman, *Phys. Rev. Lett.* 86, 4052 (2001).
422. A. A. Zakhidov, R. H. Baughman, C. Cui, I. Khairullin, M. Shkunov, and V. Z. Vardeny, in "Abstracts of the 4th International Conference on Nanostructured Materials," Nano98, Stockholm, 1998.
423. A. A. Zakhidov, R. H. Baughman, Z. Iqbal, C. Cui, I. Khayrullin, S. O. Dantas, J. Marti, and V. G. Ralchenko, *Science* 282, 897 (1998).
424. R. H. Baughman, A. A. Zakhidov, C. Cui, I. Khairullin, L. M. Liu, Z. Iqbal, S. O. Dantas, and V. G. Ralchenko, in "Abstracts of the 4th International Conference on Nanostructured Materials," Nano98, Stockholm, 1998.
425. K. Bush and S. John, *Phys. Rev. Lett.* 83, 967 (1999).
426. F. M. Aliev, V. N. Bushuev, G. N. Dulnev, I. K. Meshkovski, and A. M. Prohorov, *Dokl. Akad. Nauk* 253, 598 (1980).
427. S. M. Dolotov, M. F. Koldunov, A. A. Manenkov, G. P. Roskova, N. M. Sitnikov, N. E. Khaplanova, and T. S. Tsekhomskaya, *Kvantovaya Elektron.* 19, 1134 (1992).
428. G. N. Dulnev, V. I. Zemski, B. B. Krinetski, I. K. Meshkovski, A. M. Prokhorov, and O. M. Stelmah, *JTP Lett.* 4, 1041 (1978).
429. R. Reisfield and E. Yariv, *Tech. Univ. Wrocław. Opt. Appl.* 30, 481 (2000).
430. G. B. Altshuller, V. A. Bahanov, E. G. Dulneva, and I. K. Meshkovski, *Opt. Spektrosk.* 55, 369 (1983).
431. V. I. Zemski, Y. L. Kolesnikov, and A. V. Sechkarev, *JTP Lett.* 11, 781 (1985).
432. V. I. Zemski, Y. L. Kolesnikov, and I. K. Meshkovski, *JTP Lett.* 12, 331 (1986).
433. A. S. Eremenko, V. I. Zemski, Y. L. Kolesnikov, B. G. Malinin, I. K. Meshkovski, N. P. Savkin, V. E. Stepanov, and V. S. Shildyaev, *Opt. Spektrosk.* 61, 1114 (1986).
434. V. I. Zemskii, Y. L. Kolesnikov, and A. F. Novikov, *Proc. SPIE-Int. Soc. Opt. Eng.* 3573, 455 (1998).
435. I. Braun, G. Ihlein, F. Laeri, J. U. Noeckel, G. Schulz-Ekloff, F. Schueth, U. Vietze, O. Weiss, and D. Woehrle, *Appl. Phys. B* 70, 335 (2000).
436. V. I. Sukhanov, M. V. Khazova, A. M. Kursakova, and O. V. Andreeva, *Opt. Spectrosk.* 65, 474 (1988).
437. V. I. Sukhanov, *Opt. Appl.* 24, 13 (1994).
438. V. I. Sukhanov, *Proc. SPIE-Int. Soc. Opt. Eng.* 1238, 226 (1991).
439. O. V. Andreeva, Y. L. Korzinin, V. N. Nazarov, E. R. Gavrilyuk, and A. M. Kursakova, *Optiko-Mekhanicheskaya Promyshlennost* 64, 142 (1997).
440. V. I. Sukhanov, A. M. Kursakova, S. A. Kuchinsky, E. R. Gavrilyuk, and T. W. Zerda, *J. Sol-Gel Sci. Technol.* 8, 1111 (1997).
441. L. B. Glebov, K. V. Dukel'skii, S. K. Evstrop'ev, N. V. Nikonorov, G. T. Petrovskii, and V. S. Shashkin, *Pis'ma Zh. Tekhn. Fiz.* 16, 9 (1990).
442. V. I. Sukhanov, S. A. Kuchinskii, and O. V. Andreeva, *Opt. Spektrosk.* 81, 851 (1996).
443. C. P. Bean, M. V. Doyle, and A. G. Pincus, *Phys. Rev. Lett.* 9, 93 (1962).
444. C. P. Bean, *Rev. Mod. Phys.* 36, 31 (1964).
445. J. H. P. Watson, *J. Appl. Phys.* 37, 516 (1966).
446. J. H. P. Watson, *Phys. Rev.* 148, 223 (1966).
447. J. H. P. Watson, *Phys. Rev. B* 2, 1282 (1970).
448. P. G. de Gennes, *Phys. Kondens. Mater.* 3, 79 (1964).
449. K. Maki, *Physics* 1, 21 (1964).
450. W. L. McMillan, *Phys. Rev.* 167, 331 (1968).
451. J. W. Garland, K. H. Benneman, and F. M. Mueller, *Phys. Rev. Lett.* 21, 1315 (1968).
452. M. Strongin, O. F. Kamerer, J. E. Crow, R. D. Parks, D. H. Douglass, and M. A. Jensen, *Phys. Rev. Lett.* 21, 1320 (1968).
453. V. N. Bogomolov and N. A. Klushin, *Fiz. Tverdogo Tela* 15, 514 (1973).

454. V. N. Bogomolov, N. A. Klushin, and P. P. Seregin, *Fiz. Tverdogo Tela* 14, 2000 (1972).
455. B. Abeles, R. W. Cohen, and W. R. Stowell, *Phys. Rev. Lett.* 18, 902 (1967).
456. J. H. P. Watson, *Phys. Rev. B* 5, 879 (1972).
457. A. M. Clogston, *Phys. Rev. Lett.* 9, 266 (1962).
458. B. S. Chandrasekhar, *Appl. Phys. Lett.* 1, 7 (1964).
459. N. R. Werthamer, E. Helfand, and P. C. hohenberg, *Phys. Rev.* 147, 295 (1966).
460. L. Bosio, R. Cortes, A. Degrain, and I. Epelboin, *C. R. Acad. Sci.* 264, 1592 (1967).
461. L. Bosio, A. Defrain, J. Keyston, and J. C. Vallier, *C. R. Acad. Sci.* 261, 5431 (1965).
462. E. V. Charnaya, Y. A. Kumzerov, C. Tien, and C. S. Wur, *Solid State Commun.* 94, 635 (1995).
463. E. V. Charnaya, C. Tien, C. S. Wur, and Y. A. Kumzerov, *Physica C* 273, 91 (1996).
464. E. V. Charnaya, C. Tien, C. S. Wur, and Y. A. Kumzerov, *Physica C* 269, 313 (1996).
465. C. Tien, C. S. Wur, K. J. Lin, J. S. Hwang, E. V. Charnaya, and Y. A. Kumzerov, *Phys. Rev. B: Condens. Matter* 54, 11880 (1996).
466. E. V. Charnaya, C. Tien, K. J. Lin, C. S. Wur, and Y. A. Kumzerov, *Phys. Rev. B* 58, 467 (1998).
467. M. J. Graf, T. E. Huber, and C. A. Huber, *Phys. Rev. B* 45, 3133 (1992).
468. F. Dong, M. J. Graf, T. E. Huber, and C. I. Huber, *Acad. Sci. Czech. Republic* 46, 2367 (1996).
469. F. Dong, M. J. Graf, T. E. Huber, and C. I. Huber, *Solid State Commun.* 101, 929 (1997).
470. S. G. Romanov and D. V. Shamshur, *Sov. Phys. Solid State* 42, 581 (2000).
471. K. Wan, A. K. Jain, and J. E. Lukens, *Appl. Phys. Lett.* 54, 1805 (1989).
472. S. Pagano, R. Monaco, and G. Constablile, *IEEE Mag.* 25, 1805 (1989).
473. K. K. Likharev, "Dynamics of Josephson Junctions and Circuits." Gordon and Breach Science, Amsterdam, 1986.
474. V. N. Bogomolov and V. K. Krivosheev, *Sov. Phys. Solid State* 14, 1238 (1972).
475. V. N. Bogomolov, N. A. Klushin, Y. A. Kumzerov, and P. A. Cheremnikh, *Sov. Phys. Solid State* 18, 1476 (1976).
476. L. G. Aslamazov and A. I. Larkin, *Sov. Phys. Solid State* 10, 875 (1968).
477. P. C. Hohenberg, *Phys. Rev.* 158, 383 (1967).
478. J. S. Langer and V. Ambegaokar, *Phys. Rev.* 164, 498 (1967).
479. D. E. McCumber and B. I. Halperin, *Phys. Rev. B* 1, 1054 (1970).
480. A. I. Larkin and Y. N. Ovchinnikov, *J. Low Temp. Phys.* 10, 407 (1973).
481. W. J. Skocpol and M. Tinkham, *Rep. Prog. Phys.* 38, 1049 (1975).
482. W. J. Skocpol, M. R. Beasly, and M. Tinkham, *J. Low Temp. Phys.* 16, 145 (1974).
483. B. I. Ivlev and N. B. Kopnin, *Sov. Phys. Usp.* 142, 435 (1984).
484. R. L. Green, C. N. King, R. B. Zubeck, and J. J. Hauser, *Phys. Rev. B* 6, 3297 (1972).
485. A. M. Goldman, J. C. Solinsky, and T. J. Magee, *J. Low Temp. Phys.* 20, 339 (1975).
486. V. N. Prigodin, *JETP* 78, 1272 (1980).
487. A. A. Abrikosov and L. P. Gorkov, *JETP* 42, 1088 (1962).
488. N. F. Mott and W. D. Twose, *Adv. Phys.* 10, 107 (1961).
489. R. Landauer, *Philos. Mag.* 21, 863 (1970).
490. D. J. Thouless, *Phys. Rev. Lett.* 39, 167 (1977).
491. D. J. Thouless, *Solid State Commun.* 34, 683 (1980).
492. N. Giordano, W. Gilson, and D. E. Prober, *Phys. Rev. Lett.* 43, 725 (1979).
493. J. T. Masden and N. Giordano, *Phys. Rev. B* 31, 6395 (1985).
494. N. Giordano, *Phys. Rev. B* 22, 5635 (1980).
495. D. E. Beutler and N. Giordano, *Phys. Rev. B* 38, 8 (1988).
496. V. N. Bogomolov, E. V. Kolla, and Y. A. Kumzerov, *Solid State Commun.* 46, 383 (1983).
497. K. Liu, C. L. Chien, and P. C. Searson, *Phys. Rev. B* 58, R14681 (1998).
498. Z. B. Zhang, J. Y. Ying, and M. S. Dresselhaus, *J. Mater. Res.* 13, 1745 (1998).
499. Z. Zhang, X. Sun, M. S. Dresselhaus, J. Y. Ying, and J. Heremans, *Phys. Rev. B* 61, 4850 (2000).
500. J. Heremens, C. M. Trush, Y. M. Lin, S. Cronin, Z. Zhang, M. S. Dresselhaus, and J. S. Mansfield, *Phys. Rev. B* 61, 2921 (2000).
501. J. Heremans, C. M. Trush, Z. Zhang, X. Sun, M. S. Dresselhaus, J. Y. Ying, and D. T. Morelli, *Phys. Rev. B* 58, R10091 (1998).
502. T. E. Huber and M. J. Graf, *Phys. Rev. B* 60, 16880 (1999).
503. D. Al-Mawlawi, C. Z. Liu, and M. Moskovits, *J. Mater. Res.* 9, 1014 (1994).
504. D. Routkevich, A. A. Tager, J. Haruyama, D. Almawlawi, M. Moskovits, and J. M. Xu, *IEEE Trans. Electron Devices* 43, 1646 (1996).
505. D. N. Davydov, J. Haruyama, D. Routkevich, B. W. Statt, D. Ellis, M. Moskovits, and J. M. Xu, *Phys. Rev. B* 57, 13550 (1998).
506. L. D. Hicks and M. S. Dresselhaus, *Phys. Rev. B* 47, 16631 (1993).
507. D. A. Broido and T. L. Reinecke, *Appl. Phys. Lett.* 67, 100 (1995).
508. Y.-M. Lin, X. Sun, and M. S. Dresselahaus, *Phys. Rev. B* 62, 4610 (2000).
509. J. Heremans and C. M. Trush, *Phys. Rev. B* 59, 12579 (1999).
510. F. D. M. Haldane, *J. Phys. C: Solid State Phys.* 14, 2585 (1981).
511. J. Voit, *Rep. Prog. Phys.* 58, 977 (1995).
512. H. J. Schulz, *Phys. Rev. Lett.* 71, 1684 (1993).
513. C. L. Kane and M. P. A. Fisher, *Phys. Rev. B* 46, 15233 (1992).
514. C. L. Kane and M. P. A. Fisher, *Phys. Rev. Lett.* 76, 3192 (1996).
515. I. A. Romanovsky, I. V. Krive, E. N. Bogachek, and U. Landman, *Phys. Rev. B* 65, 75115 (2002).
516. S. Kawai and R. Ueda, *J. Electrochem. Soc.* 122, 32 (1975).
517. T. M. Whitney, J. S. Jiang, P. Searson, and C. Chien, *Science* 261, 1316 (1993).
518. L. Sun, P. C. Searson, and C. L. Chien, *Appl. Phys. Lett.* 74, 2803 (1999).
519. L. Sun, P. C. Searson, and C. L. Chien, *Phys. Rev. B* 61, R6463 (2000).
520. J. Meier, B. Doudin, and J. P. Ansermet, *J. Appl. Phys.* 79, 6010 (1996).
521. J.-E. Wegrowe, D. Kelly, A. Frank, S. E. Gilbert, and J. P. Ansermet, *Phys. Rev. Lett.* 82, 3681 (1999).
522. Y. Jaccard, P. Guittienne, D. Kelly, J. E. Wegrowe, and J. P. Ansermet, *Phys. Rev. B* 62, 1141 (2000).
523. M. Lederman, R. OBarr, and S. Schultz, *IEEE Trans. Magn.* 31, 3793 (1995).
524. L. Piraux, S. Dubois, E. Ferain, R. Legras, K. Ounadjela, J. M. George, J. L. Maurice, and A. Fert, *J. Magn. Magn. Mater.* 165, 352 (1997).
525. S. Dubois, J. Colin, J. L. Duvail, and L. Piraux, *Phys. Rev. B* 61, 14315 (2000).
526. U. Ebels, J. L. Duvail, P. E. Wigen, L. Piraux, L. D. Buda, and K. Ounadjela, *Phys. Rev. B* 64, 144421 (2001).
527. L. Piraux, J. M. George, J. F. Despres, C. Leroy, E. Ferain, and R. Legras, *Appl. Phys. Lett.* 65, 2484 (1994).
528. J. J. Thomson, "Application of Dynamics to Physics and Chemistry." Dawsons of Pall Mall, London, 1968.
529. H. W. Foot and B. J. Saxton, *J. Am. Chem. Soc.* 38, 588 (1916).
530. H. W. Foot and B. Saxton, *J. Am. Chem. Soc.* 39, 1103 (1917).
531. A. S. J. Coolidge, *J. Am. Chem. Soc.* 46, 596 (1924).
532. I. D. Jones and R. A. Cortner, *J. Phys. Chem.* 36, 387 (1932).
533. W. A. Patrick and W. E. Land, *J. Phys. Chem.* 38, 1201 (1934).
534. W. O. Milligan and H. Rachford, *J. Am. Chem. Soc.* 70, 2922 (1948).
535. M. J. Brown and A. G. Foster, *Nature* 169, 37 (1952).
536. W. A. Patrick and W. A. Kemper, *J. Phys. Chem.* 42, 369 (1938).

537. H. K. Christenson, *J. Phys.: Condens. Matter* 13, R95 (2001).
538. K. M. Unruh, T. E. Huber, and C. A. Huber, *Phys. Rev. B* 48, 9021 (1993).
539. R. W. Batchelor and A. G. Foster, *Trans. Faraday. Soc.* 40, 300 (1944).
540. J. L. Tell and H. J. Maris, *Phys. Rev. B* 28, 5122 (1983).
541. E. D. Adams, K. Uhlig, Y.-H. Tang, and G. E. Haas, *Phys. Rev. Lett.* 52, 2249 (1984).
542. D. Wallacher and K. Knorr, *Phys. Rev. B* 63, 104202 (2001).
543. B. Schafer, D. Balszunat, W. Langel, and B. Asmussen, *Mol. Phys.* 89, 1057 (1996).
544. P. Huber and K. Knorr, *Phys. Rev. B* 60, 12657 (1999).
545. D. W. Brown, P. E. Sokol, and S. N. Ehrlich, *Phys. Rev. Lett.* 81, 1019 (1998).
546. D. D. Awschalom and J. Warnock, *Phys. Rev. B* 35, 6779 (1987).
547. P. E. Sokol, W. J. Ma, K. W. Herwig, W. M. Snow, Y. Wang, J. Koplik, and J. Banavar, *Appl. Phys. Lett.* 61, 777 (1992).
548. V. N. Bogomolov and A. I. Zadorozhnii, *Sov. Phys. Solid State* 17, 1078 (1975).
549. V. N. Bogomolov, E. V. Kolla, and Y. A. Kumzerov, *JETP Lett.* 41, 34 (1985).
550. E. V. Charnaya, C. Tien, K. J. Lin, and Y. A. Kumzerov, *J. Phys.: Condens. Matter* 32, 7273 (1998).
551. C. Tien, C. S. Wur, K. J. Lin, J. S. Wang, E. V. Charnaya, and Y. A. Kumzerov, *Phys. Rev. B* 54, 11880 (1996).
552. E. Shabanova, E. V. Charnaya, K. Schaumburg, and Y. A. Kumzerov, *J. Magn. Reson., Ser. A* 122, 67 (1996).
553. E. V. Charnaya, T. Loeser, D. Michel, C. Tien, D. Yaskov, and Y. A. Kumzerov, *Phys. Rev. Lett.* 88, 097602 (2002).
554. B. F. Borisov, E. V. Charnaya, T. Loeser, D. Michel, C. Tien, and Yu. A. Kumzerov, *J. Phys.: Condens. Matter* 11, 10259 (1999).
555. J. M. Dereppe, B. F. Borisov, E. V. Charnaya, A. V. Shelyapin, N. M. Nassar, and Y. A. Kumzerov, *Phys. Solid State* 42, 193 (2000).
556. B. F. Borisov, E. V. Charnaya, Y. A. Kumzerov, A. K. Radzhabov, and A. V. Shelyapin, *Solid State Commum.* 92, 531 (1994).
557. T. Takamuku, M. Yamagami, H. Wakita, Y. Masuda, and T. Yamaguchi, *J. Phys. Chem. B* 101, 5730 (1997).
558. D. C. Steytler, J. C. Dore, and C. J. Wright, *J. Phys. Chem.* 87, 2458 (1983).
559. M.-C. Bellisent-Funel, J. Lal, and L. Bosio, *J. Chem. Phys.* 98, 4246 (1993).
560. K. Overloop and L. V. Gerven, *J. Magn. Reson., Ser. A* 101, 179 (1993).
561. Y. Harama, T. Takahashi, M. Hino, and T. Sato, *J. Colloid Interface Sci.* 184, 349 (1996).
562. J. M. Baker, J. C. Dore, and P. Behrens, *J. Phys. Chem. B* 101, 6226 (1997).
563. K. Morishige and K. Nobuoka, *J. Chem. Phys.* 107, 6965 (1997).
564. J. C. Dore, M. Dunn, T. Hasebe, and J. H. Strange, *Colloids Surf.* 36, 199 (1989).
565. R. Mu and V. M. Malhotra, *Phys. Rev. B* 44, 4296 (1991).
566. H. F. Booth and J. H. Strange, *Mol. Phys.* 93, 263 (1998).
567. R. Mu, Y. Xue, D. O. Henderson, and D. O. Frazier, *Phys. Rev. B* 53, 6041 (1996).
568. M. Sliwinska-Bartkowiak, J. Gras, R. Sikorski, R. Radhakrishnan, L. Gelb, and K. E. Gubbins, *Langmuir* 15, 6060 (1999).
569. T. Takei, Y. Onoda, M. Fuji, T. Watanabe, and M. Chikazawa, *Thermochim. Acta* 352–353, 199 (2000).
570. C. L. Jackson and G. B. McKenna, *J. Chem. Phys.* 93, 9002 (1990).
571. D. Morineau, G. Dosseh, and C. Alba-Simonesco, *Philos. Mag. B* 79, 1847 (1999).
572. J.-M. Zanotti, M.-C. Bellisent-Funel, and S.-H. Chen, *Phys. Rev. B* 59, 3084 (1999).
573. V. Crupi, D. Majolino, P. Migliardo, and V. Venuti, *Phys. Rev. E* 54, 5377 (1996).
574. R. Pelster, *Phys. Rev. B* 59, 9214 (1999).
575. M. Arndt, R. Stannarius, W. Gorbatschow, and F. Kremer, *Phys. Rev. E* 54, 5377 (1996).
576. W. Gorbatschow, M. Arndt, R. Stannarius, and F. Kremer, *Europhys. Lett.* 35, 719 (1996).
577. J. Schuller, R. Richert, and E. Fischer, *Phys. Rev. B* 52, 15232 (1995).
578. J. Schuller, Y. B. Mel'nichenko, R. Richert, and E. Fischer, *Phys. Rev. Lett.* 73, 2224 (1994).
579. Y. B. Mel'nichenko, J. Schuller, R. Richert, B. Ewen, and C.-K. Loong, *J. Chem. Phys.* 103, 2016 (1995).
580. F. Kremer, *J. Non-Cryst. Solids* 305, 1 (2002).
581. L. Fu and L. Resca, *Phys. Rev. B* 47, 16194 (1993).
582. M. Silwinska-Bartkowiak, G. Dudziak, R. Sikorski, and R. Gras, *J. Chem. Phys.* 114, 950 (2001).
583. D. Daoukaki, G. Barut, R. Pelster, G. Nimtz, A. Kyritsis, and P. Pissis, *Phys. Rev. B* 58, 5336 (1998).
584. G. Barut, P. Pissis, R. Pelster, and G. Nimtz, *Phys. Rev. Lett.* 80, 3543 (1998).
585. H. Vogel, *Phys. Z.* 22, 645 (1921).
586. G. S. Fulcher, *J. Am. Chem. Soc.* 8, 789 (1925).
587. F. Brouers and A. Ramsamugh, *J. Phys. C* 21, 1839 (1988).
588. P. Wong, J. Koplik, and J. P. Tomanic, *Phys. Rev. B* 30, 6606 (1984).
589. F. Claro and F. Brouers, *Phys. Rev. B* 40, 3261 (1989).
590. S. V. Pankova, V. V. Poborchii, and V. G. Solovev, *J. Phys.: Condens. Matter* 8, L203 (1996).
591. A. V. Fokin, Y. A. Kumzerov, N. M. Okuneva, A. A. Naberezhnov, S. B. Vakhrushev, I. V. Golosovsky, and A. Kurbakov, *Phys. Rev. Lett.* 89, 175503 (2002).
592. E. V. Colla, A. V. Fokin, E. Y. Koroleva, Y. A. Kumzerov, S. B. Vakhrushev, and B. N. Savenko, *Nanostruct. Mater.* 12, 963 (1999).
593. K. Uchino, E. Sadanaga, and T. Hirose, *J. Am. Ceram. Soc.* 72, 1555 (1989).
594. Y. Guo, K. H. Langley, and F. E. Karasz, *Phys. Rev. B* 50, 3400 (1994).
595. W. D. Dozier, J. M. Drake, and J. Klafter, *Phys. Rev. Lett.* 56, 197 (1986).
596. M. T. Bishop, K. H. Langley, and F. E. Karasz, *Phys. Rev. Lett.* 57, 1741 (1986).
597. S. Mitra, R. Mukhopadhyay, K. T. Pillai, and V. N. Vaidya, *Solid State Commun.* 105, 719 (1998).
598. L. M. Pismen, *Chem. Eng. Sci.* 29, 1227 (1974).
599. K. Hahn, J. Kärger, and V. Kukla, *Phys. Rev. Lett.* 76, 2762 (1996).
600. J. Karger and D. Ruthven, "Diffusion in Zeolites and Other Microporous Materials." Wiley, New York, 1992.
601. S. Stapf and R. Kimmich, *Chem. Phys. Lett.* 275, 261 (1997).
602. D. W. Aksnes and L. Gjerdaker, *J. Mol. Struct.* 475, 27 (1999).
603. L. Gjerdaker, D. W. Aksnes, G. H. Sorland, and M. Stocker, *Microporous Mesoporous Mater.* 42, 89 (2001).
604. A. Mertelj and M. Copic, *Phys. Rev. E* 61, 1622 (2000).
605. P. G. G. Gennes and J. Prost, "The Physics of Liquid Crystals." Oxford University Press, Oxford, 1993.
606. E. B. Priestley, P. J. Wojtowitz, and P. Sheng, "Introduction to Liquid Crystals." Plenum, New York, 1974.
607. M. Noboli and G. Duran, *Phys. Rev. A* 46, R6174 (1992).
608. G. P. Crawford, D. W. Allender, and J. W. Doane, *Phys. Rev. A* 45, 8693 (1992).
609. S. Kralj and S. Zumer, *Phys. Rev. E* 51, 2157 (1995).
610. D. E. Feldman, *Phys. Rev. Lett.* 84, 4886 (2000).
611. A. Maritan, M. Cieplak, T. Bellini, and J. R. Banavar, *Phys. Rev. Lett.* 72, 4113 (1994).
612. X. Wu, W. I. Goldberg, M. X. Liu, and J. Z. Xue, *Phys. Rev. Lett.* 69, 470 (1992).
613. T. Bellini, N. A. Clark, C. D. Muzny, L. Wu, C. W. Garland, D. W. Schaefer, and B. J. Oliver, *Phys. Rev. Lett.* 69, 788 (1992).

614. S. Kralj, A. Zidanšek, G. Lahajnar, I. Muševic, S. Zumer, R. Blinc, and M. M. Pintar, *Phys. Rev. E* 53, 3629 (1996).
615. H. Xu, J. K. Vij, A. Rappaport, and N. A. Clark, *Phys. Rev. Lett.* 79, 249 (1997).
616. L. Wu, B. Zhou, C. W. Garland, T. Bellini, and D. W. Schaefer, *Phys. Rev. E* 51, 2157 (1995).
617. G. S. Iannacchione and D. Finotello, *Phys. Rev. E* 50, 4780 (1994).
618. G. P. Sinha and F. M. Aliev, *Phys. Rev. E* 58, 2001 (1998).
619. G. Schwalb and F. W. Deeg, *Phys. Rev. Lett.* 74, 1383 (1995).
620. K. R. Atkins, H. Seki, and E. U. Condon, *Phys. Rev.* 102, 582 (1956).
621. P. A. Crowell, F. W. V. Keuls, and J. D. Reppy, *Phys. Rev. Lett.* 75, 1106 (1995).
622. D. Finotello, K. A. Gillis, A. Wong, and M. H. W. Chan, *Phys. Rev. Lett.* 61, 1954 (1988).
623. D. J. Bishop, J. E. Berthold, J. M. Parpia, and J. D. Reppy, *Phys. Rev. B* 24, 5047 (1981).
624. M. H. W. Chan, K. I. Blum, S. Q. Murphy, G. K. S. Wong, and J. D. Reppy, *Phys. Rev. Lett.* 61, 1950 (1988).
625. C. R. Anderson, K. H. Andersen, W. G. Stirling, P. E. Sokol, and R. M. Dimeo, *Phys. Rev. B* 102, 582 (2002).
626. P. E. Sokol, R. M. Dimeo, D. W. Brown, C. R. Anderson, W. G. Stirling, M. A. Adams, S. H. Lee, C. Ruitser, and S. Komarneni, *Physica B* 241–243, 929 (1998).
627. N. Mulders, R. Mehrotra, L. S. Goldner, and G. Ahlers, *Phys. Rev. Lett.* 67, 695 (1991).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanostructuring at Surfaces Using Proteins

Michael Niederweis

Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany

Stefan H. Bossmann

Universität Karlsruhe, Karlsruhe, Germany

CONTENTS

1. INTRODUCTION

The information age can be characterized by its ever-increasing demand of ultrafast computer systems, which resulted in the expectation of an exponential success in miniaturization, which is known as Moore's "law" [1]. Although more conventional techniques, such as (deep) UV-patterning [2], X-ray proximity [3], or ion-beam projection lithography [4] are being optimized and installed in industrial development and production units, it is apparent that the successful design of nanoscale systems will require entirely new materials, processes, and device architectures.

In order to meet this challenge, a wide range of methods for the nanostructuring of (mostly metal-) surfaces [5] was developed during the last two decades. This research endeavor is particularly dedicated to the development of advanced hardware. However, other commercial applications of (bio-)nanosystems are of increasing importance as well; for instance bio- and medical sensing systems [6] and optical devices [7]. The most advanced technologies today are commonly described as electrochemical nanostructuring (by tip-induced metal deposition or oxidation) [8] and atomic force microscope (AFM) lithography [9]. Electrochemical methods are based on the technology of the scanning tunneling microscope (STM) [10]. However, until now the most significant disadvantage of electrochemical nanostructuring and AFM Lithography is that nanostructuring has to be performed atom-by-atom or fragment-by-fragment, respectively. Consequently, the defined nanostructuring of surfaces has been time-consuming and expensive. Therefore, the generation of well-defined nanostructures by protein or macromolecule deposition on two-dimensional surfaces would represent a considerable advantage compared to the technologies previously mentioned. Especially the use of proteins has many advantages [11]. Proteins are, in comparison to synthetic polymers, much more "intelligent" materials, because they are able to interact site-specific with each other as well as with possible ligand molecules or substrates [12]. Furthermore, proteins permit the harvest of the results from very many of years of evolution for nanotechnical purposes [13]. Even better, the genetic technology of site-directed mutagenesis permits the introduction of specific features for tailor-made nanotechnological application into the protein-based assemblies.

1.1. Design and Future Applications of Quantum Dots

Recent developments in nanotechnology offer applications in very different areas, such as electron tunneling (superconduction) at feasible temperatures [14], (photo) electro-catalysis [15], the development of quantum computers [16], as well as novel photoelectrochemical devices, using the effect called "plasmon-resonance" of an array of nanoparticles or -wires [7]. To date, the most developed nanosystems for quantum applications consisted of

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (851–867)

doped carbon-nanotubes [17] and doped zeolites [18]. However, both template materials lead inevitably to at least some degree of electrical connection between the embedded nanoparticles. Therefore, possible quantum effects could be cloaked by unwanted electrical connection of the individual quantum dots. In this case, we cannot expect "real quantum effects" when these materials are used for advanced electrocatalysis.

The use of a stable protein as nanotemplate for the stabilization of nanoparticles and -wires would provide an electrically insulating matrix for the consequent electrosynthesis [19] of metal-nanoparticles or nanowires. Furthermore, at least some proteins possess the ability of self-organization under *in-vitro* conditions. A straightforward approach to the generation of stabilized nanoparticles consists of the use of nanotemplates, which act as "hosts" for the nanoparticles [20]. These nanoscopic particles are generated by reduction of metal salts by chemical processes [21].

Protein-stabilized nanoparticles represent an ideal quantum-dot system, having the big advantage that the exterior of the synthesized Au, Ni, or Pt-nanoparticles or -nanowires will be embedded by the protein, serving as an ideal insulator. Since only the openings of the metal-filled nanochannels or nanopores will be open to the supernatent buffer or the substrate, real quantum effects could occur. Besides quantum effects, protein-based nanodots may permit the observation of very large electrical field effects, since the electrical field at the top of a metal-doped protein channel will easily exceed 10^6 Vcm^{-1}.

1.2. S-Layer Proteins and Porins

Two groups of bacterial proteins are of a special interest for the design of future nanotechnology: the S-layer proteins and the porins. S-layer proteins form two-dimensional crystalline surface arrays and have been identified in hundreds of different species from nearly every taxonomic group of walled eubacteria and are also present in many archeabacteria [22]. S-layer proteins from numerous species were isolated and characterized and are, in principle, available for technical applications. Porins are channel proteins in bacterial outer membranes and allow the diffusion of small and hydrophilic compounds [23–26]. In contrast to the S-layer proteins, various porins have been intensively studied with respect to their biophysical properties [27–29]. The mechanisms and the triggering factors of the channel-opening are of especially great interest because of possible applications of ion- and diffusive channels [30–34]. Furthermore, the crystal structures of some porins of Gram-negative bacteria have been solved and the diffusion of ions through porin channels have been computer-simulated [35].

Channel porins from Gram-positive bacteria and especially from Mycobacteria [36] are of special interest with respect to their application in nanotechnology because of their unique channel lengths [37].

Since the structures (and especially also the functions) of bacterial porins have been optimized in evolution, we might as well use these highly versatile biological structures directly, instead of mimicking their structures and working principles in chemical membrane and surface models. Nevertheless, it must be noted that evolution was not directed toward any applications in nanotechnology, but rather headed for an optimized influx of nutrients while providing maximal protection at the same time. Therefore, we have to proceed with great care on this path. It is apparent that the chemical stability of bacterial proteins (S-layer subunits and porins) against thermal and photochemical decomposition reactions must be sufficient in order to permit a successful application in nanotechnology [38]. The porin MspA isolated from *Mycobacterium smegmatis* possesses a superior chemical stability and can, principally, be used as a chemically synthesized polymer [39]. It possesses very distinct properties and forms nanochannels (not open "pore" structures like those found within S-layer lattices) of approximately 2.5 nm in diameter and 10 nm in length [37]. Furthermore, MspA can be reconstituted in various environments just as lipid bilayers, polymers, and hydrophobic surfaces. Therefore, the formation of nanochannels cannot only be achieved within one defined environment, like in the case of pore formation within S-layer lattices, but can be achieved in virtually any hydrophobic environment. This approach is considerably more versatile with respect to nanotechnology applications. In this respect, it is very fortunate that MspA has a very high tendency to form pores in artificial environments and that these pores were, at least in the examples known to date, of comparable geometric extensions. MspA nanochannels are hydrophilic in their interior and strongly hydrophobic in their exterior. Due to the electrically insulating nature of the pore walls and their geometric dimensions, the pores formed by the MspA porin from *Mycobacterium smegmatis* represent a serious challenge to carbon nanotubes as host systems for nanotechnological applications.

This chapter is describing the most important developments in the field of nanotechnology on surfaces employing proteins. It consist of the following sections: (2) S-Layer Proteins; (3) Porins from Gram-Negative Bacteria; (4) The Porin MspA from *Mycobacterium smegmatis*.

In each section, a comprehensive description of the most striking biochemical and biophysical properties is provided, before we discuss the applications as well as the application potential of the proteins of interest in nanotechnology at surfaces and surface structures.

2. S-LAYER PROTEINS

The surface of many bacteria is covered by two-dimensional crystalline arrays of proteins or glycoprotein subunits, which are commonly referred to as "S-layers" [22, 40]. Consequently, S-layer proteins belong to the most abundant proteins in prokaryotes comprising up to 15% of the total cellular protein [22]. Although considerable knowledge has accumulated on the structure, assembly, biochemistry, and genetics of S-layers, relatively little information is available about their specific functions [41]. S-layers are thought to function as molecular sieves, protective or adhesive coatings or ion traps. Furthermore, in all bacteria that possess S-layers as exclusive cell wall components, the cell shape is determined by the (glyco)proteins [41]. Whereas the S-layer proteins from archaebacteria and Gram-negative eubacteria are not connected to their cytoplasmic or outer membranes, respectively, the S-layer

proteins of at least some Gram-positive eu- and archaebacteria are indeed chemically linked [42]. As it can be seen from the examples provided in Figure 1, considerable variations in structure, complexity, and lattice symmetry of S-layer proteins exist. However, many S-layers are similar with respect to their chemical composition, formation principles, biophysical properties, and molecular size distribution.

The thickness of S-layers ranges from 5–25 nm [40]. The outer surface appears often to be smooth, whereas the inner surface is more corrugated [43]. Structural investigations of S-layers *in vivo* and after recrystallization on various substrates were performed using a combination of very different techniques, such as metal-shadowing, negative staining, ultrathin-sectioning, freeze-etching, scanning probe microscopy experiments, and underwater atomic force microscopy [22]. Structural information was obtained down to the subnanometer range. Consequently, the formation principles of two-dimensional S-layer protein arrays and their self-repeating structures are known [44]. S-layer lattices, which form an array of protein or glycoprotein subunits, possess either oblique (p1, p2), square (p4), or hexagonal (p3, p6) symmetry (see Fig. 2). The symmetry class in which the S-layer proteins crystallize is strongly dependent on the bacterial species, but not on the experimental conditions of the crystallization process. The experimental conditions only determine whether the formation of well-ordered lattices at surfaces does occur or not. One of the results from the structure elucidation of S-layers is that they form highly porous mesh works with a porosity in the range between 30–70%. The assemblies of identical protein or glycoprotein subunits possess masses in the range of 35,000 to 230,000 Dalton and exhibit pores of identical size and morphology. Many S-layers feature two or even more distinct classes of pores in the range of 2–8 nm. The subunits of most S-layers interact with each other during the process of S-layer formation through noncovalent forces including hydrophobic interactions, hydrogen bonds, ionic bonds, and the interaction of polar groups [40].

The chemical composition of S-layer (glyco)proteins from all phylogenetic branches was found to be surprisingly similar [40]. The typical S-layer is composed of an acidic protein or glycoprotein, possessing an isoelectric point in the region from pH 3 to 5. Since the isoelectric point of a protein results from the superposition of the isoelectric points of those amino acids, which are exposed to the aqueous electrolyte, it is very interesting to take a look at the chemical composition. Typical S-layer proteins have high amounts of glutamic and aspartic acid. Together, they account for 15 mol%. The lysine content of S-layer proteins is in the range of approximately 10 mol%. Thus, ionic amino acids make up about a quarter of the total composition of S-layer proteins. This finding indicates clearly the importance of ionic interactions during the formation of the S-layer lattice structure. It is noteworthy that the content of sulfur-containing amino acids in the S-layer proteins is close to zero. Interestingly, the main fraction of amino acids of S-layer proteins is hydrophobic in nature. It is even more interesting that hydrophobic and hydrophilic amino acids do not form extended clusters, but alternating segments with a more hydrophilic region at the terminal N-end. Information regarding the secondary structures of S-layer

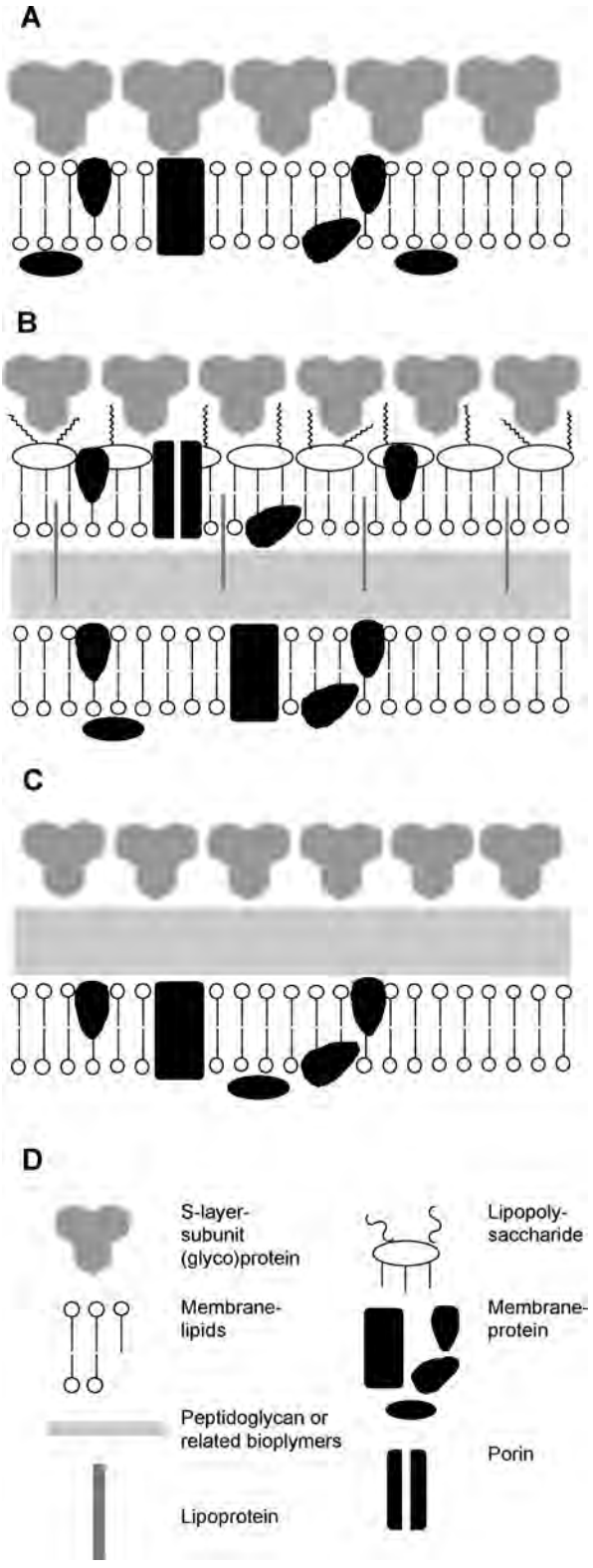


Figure 1. The structures of bacterial cell walls. (A) Archaebacteria; (B) gram-negative eubacteria; (C) gram-positive archae- and eubacteria [22]; and (D) legend.

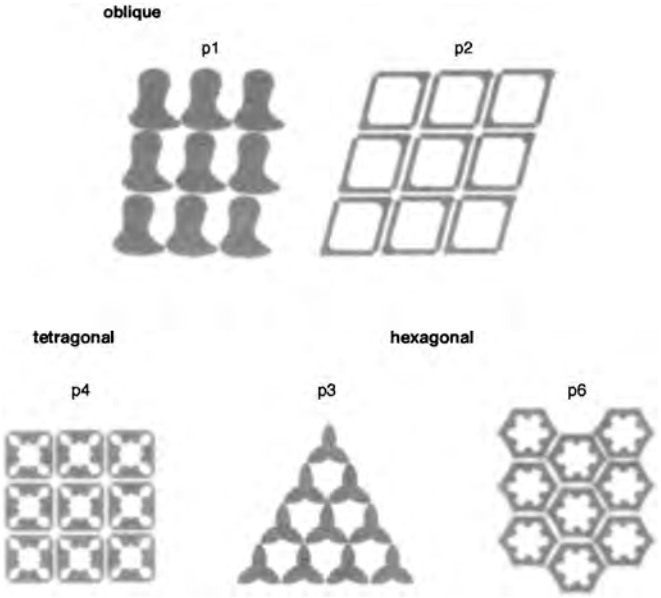


Figure 2. Symmetries of S-layer lattices (schematic representations).

proteins was obtained by using circular dichroism measurements, FT-IR-spectroscopy, and especially from secondary structure prediction based on the available protein sequence data [45]. In a typical S-layer, up to 70% of the proteins form α-helices and 20% form β-sheets. Aperiodic foldings and β-turn contents account for most of the missing fraction. In many of the S-layers, the N-terminal regions, which are among the parts of the protein-sequence self-organized as α-helices, are able to recognize distinct structures in the underlying cell envelope [46]. Because of the noncovalent nature of the bonds between the subunits, which are hydrophobic as well as ionic in nature, the growth of the S-layer, which is in most of the investigated cases surrounding the whole bacterium, proceeds very rapidly. One can approximate that a coherent S-layer at the surface region of an average cell is formed by the supramolecular interaction of approximately 5×10^5 protein or glycoprotein subunits. For S-layered bacteria with a generation time of 20 min, approximately 500 copies of the S-layer subunits have to be synthesized inside the cell, translocated and integrated into the S-layer per second [22] (Fig. 3). Because of their unusual building mechanisms and dynamics, S-layers can been described as "dynamic closed surface crystals," with the intrinsic capability to assume a structure of low free energy during cell growth and also during their growth and assembling processes at membranes and artificial surfaces.

2.1. Applications of S-Layer Proteins at Surfaces

The physicochemical properties and especially the permeability of S-layer cell envelopes are of a profound interest because they offer (a) an insight into the function principles of the bacterial transport mechanisms, and (b) the possibility of designing permeability filters for nano- and biotechnological applications [22, 47].

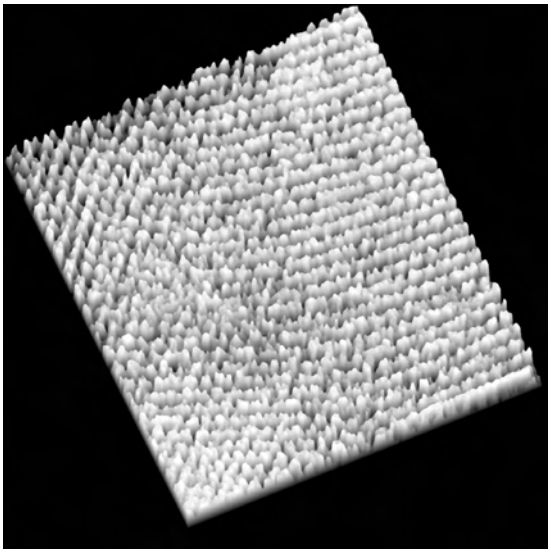


Figure 3. Oblique S-layer lattice of *Bacillus stearothermophilus* Adapted with permission, from [22], M. Sara and U. B. Sleytr, *Prog. Biophys. Molec. Biol. 65*, 83 (1996). © 1996, Elsevier Science.

2.2. S-Layer Proteins as Permeability Barriers

S-layer proteins from mesophilic and thermophilic *Bacillaceae* were studied in detail. Opposite to the underlying peptidoglycan-containing "wall skeleton," no net negative surface charge was determined. The charge density found in the native S-layer from *B. sphaericus* was 1.6 carboxylic acid functions per nm^2. However, the free carboxylic acid functions are neutralized by free amino groups of lysine. The presence of an approximately equimolar amount of amino groups and carboxylic acid groups in the pore areas of many *Bacillaceae*, prevented the adsorption of charged macromolecules inside the pores. Therefore S-layers from *Bacillaceae* can be used as suitable antifouling coatings. Furthermore, permeability studies performed on the same S-layers, demonstrated that, in spite of the considerable differences in the lattice types, the molecular weight cut-off was 30,000–40,000 Da [48]. This data is in agreement with the assumption of a pore diameter of approximately 4–5 nm.

2.3. S-Layer Ultrafiltration Membrane

For the synthesis of S-layer ultrafiltration membranes (SUMs), [49] either cell wall fragments with adhering S-layer proteins or previously isolated self-assembling S-layer subunits are used. Also in this case, S-layers from *Bacillaceae* possess ideal physical and chemical properties. The fixation of S-layers at microfiltration membranes can be achieved using a pressure procedure. Then the S-layer lattices on top of either the microfiltration membrane or the cell fragment adhering to the microfiltration membrane are crosslinked with glutaraldehyde, and thereafter the Schiff bases are reduced employing sodium borohydride as reagent. Depending on the very defined structures of S-layer lattices on surfaces, the rejection curves of S-layer-SUMs are distinctly

steeper than of microfiltration membranes made of amorphous polymers. Again, a molecular weight cut-off of 30,000–40,000 was found. This is not surprising because the molecular weight rejection effect is achieved by the pore structure of the S-layer lattice.

2.4. Chemical Modification of S-Layers

The crosslinking procedure using glutaraldehyde produces eventually a net negative charge. Carbodiimide activation offers the opportunity to link carboxylic acids to a manifold of chemical reagents and groups [50]. Thus, either a strong net negative charge or a strong positive net charge can be achieved. Also the hydrophobicity of SUMs could be modulated following the same experimental approach. It is of special importance that the distinct geometry of the self-repeating S-layer lattice did not change noticeably when the free carboxylic acids are chemically modified. S-layer glycoproteins offer the same opportunities via the activation of their glycan chains either by periodate oxidation or cyanogens-bromide activation.

A recent example of this versatile strategy was the application of S-layer-coated liposomes for entrapping and binding of functional molecules: The S-layer protein from *Bacillus stearothermophilus* PV727P2 was recrystallized on positively charged unilamellar liposomes of an average diameter of 180 nm formed by dipalmitoyl phosphatidylcholine (DPPC), cholesterol, and hexadecylamine (HDA) (molar ratio: 10/5/4) [51]. The S-layer subunits possess the molecular weight of 97,000 and form a two-dimensional lattice of oblique symmetry. Crosslinking was achieved using glutaraldehyde. Then, the free carboxylic groups were activated with 1-ethyl-3-(dimethylaminopropyl)-carbodiimide (EDC). This reagent has the advantage that it can be successfully employed also in aqueous solutions. The crosslinking process could be observed employing S-layer–coated carbonic anhydrase containing liposomes (SCALs). Whereas the crosslinking reaction proceeded exclusively between adjacent S-layer subunits within the two-dimensional lattice at the outer surface of the liposomes, the enzyme carbonic anhydrase, located inside the liposomes, was not effected by the chemical treatment as it could be demonstrated using SDS-polyacrylamide gel electrophoresis (SDS-PAGE). When the crosslinking of the S-layer subunits was performed using Bis(sulfosuccinimidyl)suberate (BS), evidence for heterologous crosslinking between the S-layer protein and membrane incorporated HDA was found.

The supramolecular as well as the highly specific binding of molecules and polymers to S-layer lattices can also be achieved by using noncovalent forces. Important factors determining the nature and the selectivity of the binding are the S-layer lattice type, the geometric sizes (or the van-der-Waals radii, respectively) of the morphological units and the chemical nature, the physicochemical properties, and the distribution of binding sites on the array. An example for this binding type is the binding of polycationic ferritin (PCF) on the S-layer lattice of *B. coagulans* E38-66 [52].

The hexagonally ordered S-layer lattices from *B. stearothermophilus* PV72 and *Thermoanaerobacter thermohydrosulfuricus* L111-69 [53] and the square S-layer lattice from *B. sphaericus* CCM 2120 can bind anionic, cationic, zwitterionic, and noncharged macromolecules. Note that only the surface of *Thermoanaerobacter thermohydrosulfuricus* is glycosylated [22].

2.5. Semifluid Membranes

A simple and straightforward method for generating coherent S-layers on lipid films is the injection of isolated S-layer protein subunits in the subphase of the Langmuir–Blodgett (LB)-trough. The crystallization of the two-dimensional S-layer lattices begins at several nucleation points and proceeds underneath the preformed phospholipid or tetraether-lipid films until coalescence of the single crystal domains occurs. Generally, the orientation of S-layers, which can be crystallized on any kind of surfaces, is determined by the anisotropy of their physicochemical surface characteristics. The more hydrophilic surface, which possesses net negative charges, is usually directed towards the negatively charged or zwitterionic head groups of phospholipid or tetraether-lipid films. Since the direction is determined by the charged head groups embedded in the S-layer structure, the more hydrophobic "outer" face (with respect to its natural orientation adhering to the bacterial cell) is directed toward the water phase during its formation. However, ordered phases with adhering S-layers can be either turned or transferred onto organic, ceramic, or metal supports. It is noteworthy that the fluidity of the lipid films, which were previously used as self-assembled mono- or bilayers is a very critical parameter for generating coherent S-layer lattices. When the formation of S-layers adhering to lipid films is achieved, crosslinking can be performed using glutaraldehyde. The latter procedure increases the mechanical stability of the S-layers on lipid film remarkably.

In a very similar manner, so-called semifluid membranes can be generated [54]. In these cases, where free carboxylic or amine groups are left after the crosslinking procedure using glutaraldehyde, lipid molecules of the LB film can be chemically linked to the S-layer protein lattice. The linking of the S-layer protein subunits to individual lipid molecules, which interact nevertheless with all lipid molecules forming the lipid monolayer or bilayer, decreases the lateral diffusion (Fig. 4). Consequently, the fluidity of the whole membrane assembly decreases. Using standard procedures, a manifold of functional biomolecules (for instance, ion channels, molecular receptors, proton pumps, or porins) can be integrated in semifluid membranes. Depending on the structure and hydrophobicity of the biomolecules, the integration into the lipid film can be achieved either before the chemical fixation of the adhering S-layer lattice or after this procedure. Because of the flexibility of this experimental approach, this technology has a big potential to initiate a broad spectrum of developments in the fields of optical or electronic nanodevices, diagnostics in medicine, and life science engineering.

The application of S-layer–coated liposomes for entrapping and binding of functional molecules via the avidin- or streptavidin-biotin bridge demonstrated this potential [54]. Two biotin residues accessible to avidin binding were bound to every S-layer subunits within the crosslinked S-layer lattice. Using biotinylated ferritin, which serves as an excellent electron paramagnetic resonance (EPR) marker, it could

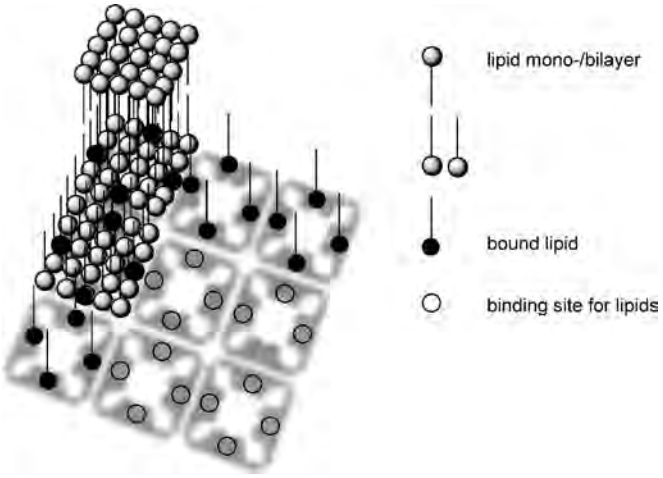


Figure 4. Schematic representation of a "semifluid" membrane, composed of an S-layer lattice possessing binding sites for lipid molecules and a lipid mono-/bilayer.

be demonstrated that a well-ordered layer of straptavidin formed on the accessible surface of the S-layer coated liposomes.

The same strategy also proved valuable when biotinylated antihuman IgG was attached via streptavidin to the biotinylated S-layer–coated liposomes (Fig. 5). The biological activity of the S-layer bound antihuman IgG was confirmed using the enzyme-linked immunosorbent assay technique (ELISA).

A similar strategy was used for the immobilization of the enzymes invertase (Mr 270,000) and glucose oxidase (Mr 150,0000) at the S-layers of *B. stearothermophilus* PV72 and *T. thermohydrosulfuricus* L111-69 [54]. The labeling density corresponded to 2–3 enzymes per hexametric unit cell of the S-layer lattice. This finding corresponded to the formation of a monomolecular enzyme layer on the surfaces of both L-layer lattices. Immobilized invertase retained approximately 70% of its enzymatic activity in the nonimmobilized state, glucose oxidase approximately 35%. When spacer molecules were used (4-amino butyric acid or 6-amino caprotic acid), the measured activity of glucose oxidase increased to about 60%. The use of spacers obviously prevented the binding of whole enzymes or partial enzyme structures within the S-layer pores and is especially advantageous for the generation of enzyme assays.

2.6. S-Layer Based Amperometric and Optical Biosensors

As already pointed out, a broad range of amperometric and optical bioanalytical sensors have been developed using suitable S-layer lattices as nanoscale immobilization matrix. The concepts of enzyme-immobilization and the manufacture of SUMs have already been discussed. For the fabrication of a single enzyme sensor, for instance a glucose sensor, glucose oxidase was bound to the surface of a S-layer ultrafiltration membrane. The electrical connection to the sensing layer was achieved by sputtering a nanoscopic gold or platinum layer onto the S-layer bound enzyme layer. The whole sensor assembly is then usually stabilized using a conventional gold or platinum electrode, by which the electrical contact to the nanolayer is achieved. In this configuration, the analyte reaches the immobilized enzyme through the pores of the S-layer (Fig. 6) [55, 56].

2.7. Molecular Nanotechnology Using S-Layers

One important and very successful approach towards the molecular nanotechnology of the future consists in the generation of so-called "self-assembled monolayers" (SAMs) at metal (especially gold)-, silicon-, or highly oriented pyrolithic graphite (HOPG) surfaces [57]. Self-assembled monolayers are formed by chemically attached lipids, which are bound to these surfaces in such a density that

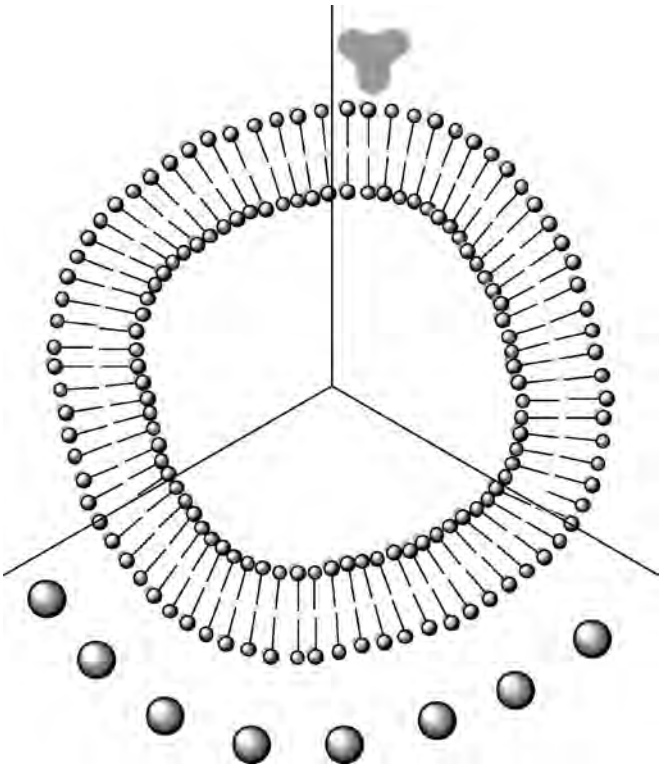


Figure 5. Schematic representation of a liposome, which is coated with an S-layer lattice. The S-layer subunits are chemically linked to functional macromolecules, such as enzymes.

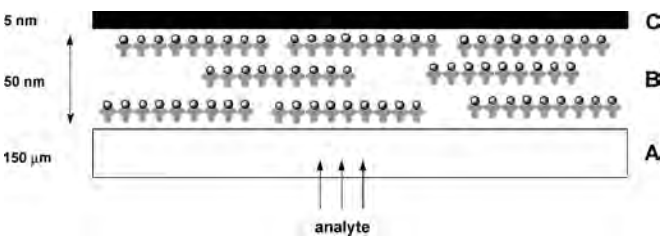


Figure 6. Schematic representation of an amperometric biosensor. (A) microfiltration membrane; (B) multiplayer, consisting of enzyme loaded S-layer fragments; (C) sputtered metal layer.

cooperative effects, just as two-dimensional condensation, occur. Therefore, SAMs are highly ordered and can serve as substrates for the binding and/or incorporation of hydrophobic substrates of all kinds. A very interesting alternative to the formation of SAMs is provided by the recrystallization of S-layer subunits of surfaces, which are suitable for nanofabrication. Silicon or gallium arsenide wafers are two examples for the successful creation of an ordered S-layer on a semiconductor material. S-layer lattices have a great potential for the development of molecular manufacturing procedures and biological nanoresists.

A first success using recrystallized S-layer lattices on semiconductor materials was achieved by the method of deep ultraviolet irradiation (DUV) [58] (Fig. 7). This method is principally related to the conventional negative photoresists, which will not be able to satisfy the hardware needs of the information age. However, it can be regarded as an example for the modern nanotechnology yet to be developed, because it uses the self-replication properties of S-layer lattices for lateral nanopatterning of wafer surfaces.

The method of deep ultraviolet irradiation employs the argon-fluorine (ArF) excimer- (or more correctly, exciplex-) laser system as a coherent light source. The current restriction in the availability of optimal mask systems only permits the generation of submicron structures. Exposure with deep ultraviolet photons, followed by the treatment with an aqueous buffer solution leads to the complete removal of the S-layer lattice located on top of the wafer material. It is very likely that the absorption of high energy photons causes chemical transformations of the amino acids, which are forming the S-layer subunit proteins. Possible photoreactions are Norrish I and II reactions [59], as well as photoinitiated oxidation reactions, which produce water-soluble groups such as carboxylic acids and alcohols. Because of this deep UV-induced chemical transformation, the ability of the S-layer subunits to form ordered lattices is greatly diminished. If the exposure time is optimized, the exposed S-layer surfaces can be completely removed by this procedure. It is noteworthy that the unexposed S-layer surface areas retain their structural and, therefore, their functional integrity.

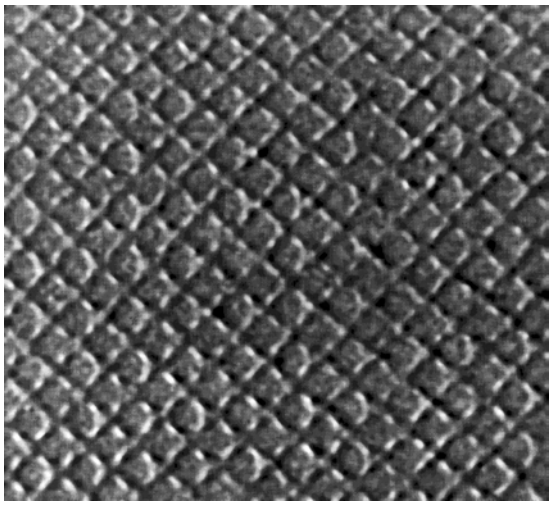


Figure 7. DUV (193 nm)-patterned S-layer coated surface (550 × 550 nm). (For the original images, see [58]).

2.8. Synthesis of Semiconductor and Metal Nanoparticles Using S-Layer Templates

As pointed out earlier, the use of organic and inorganic templates for the synthesis of semiconductor and metal nanoparticles is a very promising strategy. S-layer lattices represent ideal templates, because they represent highly ordered pore structures [57]. By choosing the right S-layer type, nanopores can be obtained. These nanopores can support either the crystallization of inorganic nanoparticles or their chemical or electrochemical reduction [54].

Inorganic superlattices of cadmium sulfide (CdS) with either oblique or square lattice symmetries were fabricated by exposing self-assembled S-layer lattices to Cd^{2+}-ion solutions. After cation-exchange, a slow reaction with hydrogen sulfide was carried out. Precipitation of CdS was confined to the nanopores only and, therefore, the resulting CdS-doped S-layer superlattices possessed the same symmetry as the S-layer lattices used as templates. Using this method, extended arrays of a very defined nanostructure, extending to 5×10^4 nm^2 could be fabricated [60].

A similar procedure resulted in the generation of a square superlattice of uniform 4–5 nm-sized gold particles with a 12.8 nm repeat distance [61]. The key steps of this procedure were the induction of thiol groups to the S-layer and the exposition of the resulting template to a tetrachloroauric acid ($H[AlCl_4]$) solution. Electron irradiation under transmission electron micrograph (TEM) conditions led to the initial formation of a grainy gold layer covering the whole surface of the S-layer. The shape of the gold particles resembled the morphology of the pore region of the S-layer. Crystalline Au(O)-particles were found. However, they did not follow a crystallographic pattern. This approach to nanopatterned metal-coated S-layers did not only work with $H[AlCl_4]$, but also employed $PdCl_2$, $NiSO_4$, $KPtCl_6$, $Pb(NO_3)_2$, and $K_3[Fe(CN)_6]$ [54]. Principally, any combination of S-layer lattice symmetries (oblique, square, or hexagonal), -lattice dimensions and pore sizes with a suitable nanomaterial (metal, metal oxide, or semiconductor particle) can be achieved.

3. PORINS FROM GRAM-NEGATIVE BACTERIA

Porins are channel proteins in bacterial outer membranes and allow the diffusion of small and hydrophilic compounds. Their properties are under intensive investigation [23–26]. Numerous biotechnological applications of porins from Gram-negative organisms were reported [62, 63] with respect to biochemical [64–66] and medicinal [67] applications. Porins from *E. coli* or other Gram-negative bacteria are stable channel-forming proteins, but they dissociate into inactive monomers above temperatures of 55–70 °C [68–70]. Therefore, their application in nanotechnology at surfaces are limited.

We want to describe one very promising application of a porin from a Gram-negative organism in more detail here, because of its prospects for future developments: Nanoreactors based on polymerized ABA-triblock copolymer vesicles have been successfully designed [71]. They

consisted of stable nanocapsules from an ambiphilic poly(2-methyloxazoline)-block-poly(dimethylsiloxane)-block-poly(2-methyloxazoline) triblock copolymer (PMOXA-PDMS-PMOXA) [72]. In perfect structural analogy to the S-layer proteins, these artificial block copolymers consist of separated hydrophilic and hydrophobic polymer segments.

Therefore, the PMOXA-PDMA-PMOXA-triblock copolymer aggregates spontaneously in dilute solutions into vesicular structures. Depending on the experimental conditions, the size of the polymer-triblock vesicles can be controlled in the range from 50–500 nm [73]. Again, in a further analogy of the S-layer lattices, the resulting nanocapsules can be further stabilized by subsequent crosslinking polymerization of the reactive end groups. The result is a "supermacromolecule," in which all the individual block copolymers are connected by covalent bonds. The polymer triblock vesicles were used to encapsulate a model enzyme β-lactamase in its interior [74]. To create a nanoreactor, the very thick triblock copolymer shell ($d = 10$ nm) has to be overcome in order to feed the enzyme with a suitable substrate. However, owing to their higher hydrophobic thickness, the triblock copolymer shells are even less permeable than conventional phospholipid bilayers. This situation is a typical case where the application of a suitable channel porin is requested, because the PMOXA-PDMA-PMOXA triblock copolymer behaves in aqueous buffer solutions principally like a higher molecular weight analogue of conventional lipids. However, the block copolymer membranes are approximately twice as thick as conventional bilayers.

The bacterial porin OmpF forms trimeric channels in the outer membrane of *E. coli* [75] and was chosen to provide a pathway through the block copolymer membrane. The high flexibility and the conformational freedom of the triblock-copolymer membrane permits its adaptation to the specific requirements of the bacterial porin [76]. If the OmpF porin can retain its specific geometric and dynamic conditions, it will retain its "natural" functions. After the spontaneous reconstitution of OmpF in the triblock copolymer shells, the encapsulated enzyme β-lactamase was successfully "fed" using the antibiotic ampillicin as model substrate, resulting in the "production" of ampicillinoic acid (Fig. 8). The latter, in contrast to ampillicin, is able to reduce iodine to iodide. This offers the opportunity to observe the reaction progress monitoring the Vis-absorption of the iodine/starch-complex. A high activity of β-lactamase was retained within the nanoreactors, which represent a perfect model for the nanoreactor technology of the future. Upscaling of a nanoreactor system is easily possible because of the very favorable diffusion conditions in the nanoscopic scale.

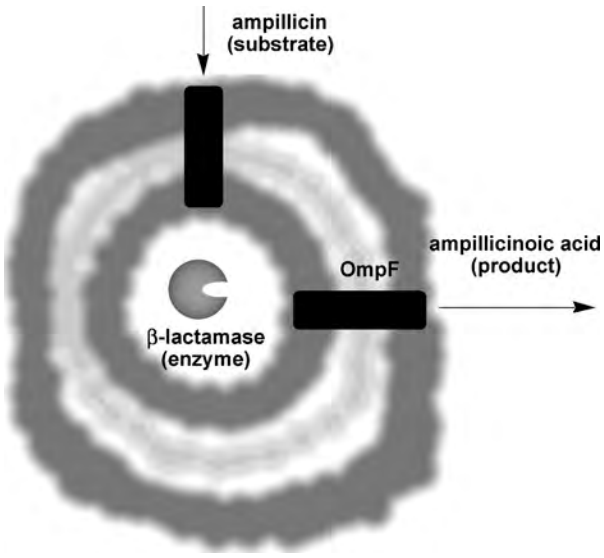


Figure 8. Scheme of a nanoreactor consisting of the crosslinked PMOXA-PDMA-PMOXA triblock copolymer, the bacterial porin OmpF, and the encapsulated enzyme β-lactamase [74].

4. THE PORIN MspA FROM *MYCOBACTERIUM SMEGMATIS*

4.1. The Mycobacterial Cell Envelope

The mycobacterial cell envelope forms an exceptionally strong permeability barrier rendering mycobacteria naturally resistant to a wide variety of antimicrobial agents. This is due to its unique structure [77]. The arrangement of the various layers of the mycobacterial cell wall is shown schematically in Figure 9. The innermost layer of the envelope is the cytoplasmic membrane of about 4 nm thickness. Outside this membrane is the so-called "cell wall skeleton," a giant macromolecule entirely surrounding the bacterial cell,

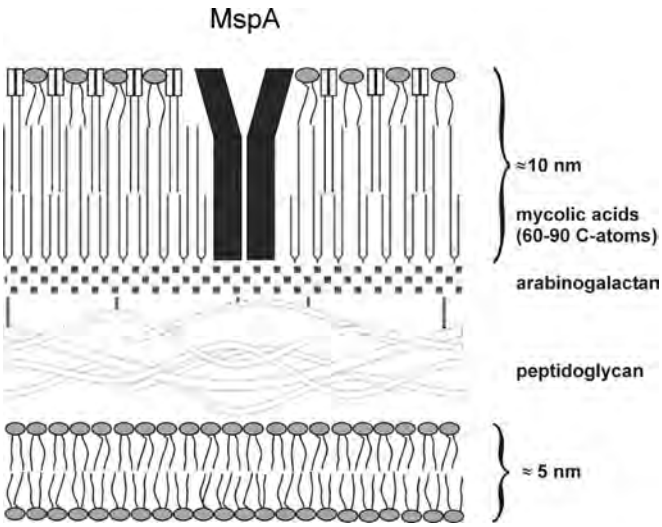


Figure 9. Structure of the mycobacterial cell envelope The inner leaflet of the outer membrane is composed of very long-chain fatty acids, the mycolic acids, which are covalently linked to the arabinogalactan-peptidoglycan copolymer The mycolic acids differ in length and modifications. The outer leaflet is formed by a great variety of extractable lipids such as trehalose-dimycolate ("cord factor"), lipooligosaccharides, sulfolipids, glycopeptidolipids, phenolic glycolipids and glycerophospholipids. It should be noted that bilayer formation of mycolic acids is only possible, when the cross-linked glycan strands and the arabinogalactan strands run in a direction perpendicular to the cytoplasmic membrane. The inner membrane is mainly is mainly composed of convential phospholipids. The thicknesses of the inner and outer membrances are derived from electron microscopic images of mycobacterial cell envelopes and are drawn to scale.

and consisting of peptidoglycan (a structure of oligosaccharides formed from disaccharide units of N-acetylglucosamine and N-glycolylmuramic acid crosslinked by short peptides), arabinogalactan (a complex branched polysaccharide), and mycolic acids (long-chain, 2-alkyl-3-hydroxy fatty acids). Associated with the cell-wall skeleton, but not covalently attached to it, are a large variety of other lipids. Substantial evidence for the organization of the mycolic acids in a second lipid bilayer in addition to the cytoplasma membrane was presented by Nikaido et al. [78]. X-ray diffraction measurements on purified mycobacterial envelopes, free of plasma membranes (<2% contamination or less), discovered a strong reflection at 4.2 Å and a weaker and more diffuse one at 4.5 Å. Such reflections are distinctive of ordered fatty acyl chains, and were interpreted as indicating the presence of highly ordered and less ordered regions, respectively. Measurements on a partly oriented sample of walls obtained by centrifuging onto a flat surface showed that the acyl chains were aligned perpendicular to the plane of the wall. By studying purified walls (walls with most of their associated lipids removed with the detergent Triton X-114 and whole cells) it has been clearly shown that the nature of the mycolic acids is the main determinant of the high-temperature phase change. Chain length is one important factor, since corynebacteria, whose mycolic acids are much shorter than those of mycobacteria, have a lower temperature phase change. Also important seems to be the configuration of the double bond or cyclopropyl group proximal to the carboxyl group of the mycolate, since a higher temperature phase change correlates with a higher proportion of *trans* configuration. It was shown for some environmental mycobacteria that they can adjust the composition of their mycolates according to temperature, presumably to achieve the required behavior of their outer permeability barrier. So *M. smegmatis* became less permeable to two lipophilic drugs when grown at higher temperatures [79, 80].

Measurements by continuous wave electron paramagnetic resonance (CW-EPR) of lipophilic probes (spin-labeled fatty acids) "dissolved" in purified cell walls or whole bacteria showed that these entered only a less ordered and more fluid region. This second region may either be interpreted as the alkyl chains of associated lipids that intercalate into that part of the mycolate monolayer where only the longer of the two alkyl chains of each mycolate is present or as the region occupied by the alkyl chains of the associated lipids forming the other (outer) half of a bilayer. The measured mobility varied with depth of insertion of the spin label into the layer (determined by the position of carbon atom in the fatty acids that was substituted with the spin label of the nitroxide type). This effect had already been observed with conventional bilayers, but the change of mobility with depth was different in the case of mycobacterial walls, underlining the unusual nature of the mycobacterial outer permeability barrier. Electron paramagnetic resonance spectra obtained with whole cells were similar to those obtained with highly purified walls, confirming that the nitroxide-labeled fatty acids entered only the outer part of the barrier [81, 82].

Detailed knowledge of the structure of the arabinogalactan of the mycobacterial envelope made it clear how a mycolate monolayer can be formed even though the mycolate residues are covalently attached to the polysaccharide. The mycolates occur as esters of terminal arabinose units on the polysaccharide, four mycolates attached to five arabinoses. The arabinosyl mycolate units are chemically bound to arabinan side chains, which are attached in turn to a galactan backbone, quite close to its attachment to the peptidoglycan. The whole polysaccharide is composed of sugars in their furanose form, which confers additional flexibility on the chain and so presumably allows the structure to accommodate itself to the close packing of the mycolate units [78, 81].

All these observations support the view of a mycobacterial outer permeability barrier as an asymmetric bilayer, with an inner leaflet of closely packed, essentially "frozen" mycolate residues and an outer leaflet of more mobile lipids. This outer membrane has unique properties: (i) it has a very low fluidity and does not melt up to 70 °C in contrast to cytoplasmic membranes of other mesophilic organisms, which begin to melt at 20 °C [82]; (ii) it is about 10 nm thick, exceeding the thickness of all other known membranes by about 2.5-fold [83]; (iii) it provides a very hydrophobic cell surface, which causes the bacteria to clump in an hydrophilic environment; and (iv) its fluidity decreases toward the periplasmatic side of the membrane in contrast to that of the outer membrane of Gram-negative bacteria [81].

4.2. MspA from *M. smegmatis* is the Prototype of a New Family of Bacterial Porins

Hydrophilic molecules enter the mycobacteria by diffusing through channel-forming proteins—the porins [82, 83]. MspA was recently identified as the first porin from *M. smegmatis* [36]. An MspA deletion mutant demonstrated that it provides the main hydrophilic pathway through the cell wall of *M. smegmatis* [84]. Electron microscopy and crosslinking experiments revealed that MspA is a tetrameric protein forming a central pore of 10 nm in length [37] (Fig. 10).

This is drastically different from the trimeric porins of Gram-negative bacteria, which have three pores per molecule and are 4 nm long [26]. These properties classify MspA as the prototype of a new family of pore proteins, since three other porins MspB, MspC, and MspD were identified in *M. smegmatis*, which differ only by a few amino acids from MspA [85].

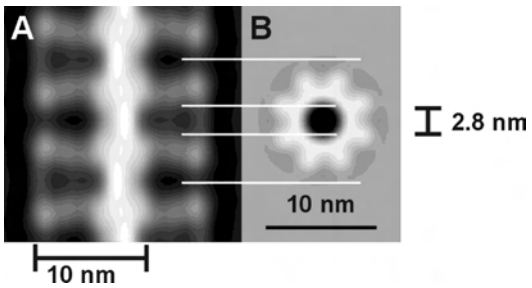


Figure 10. Structure of MspA by electron microscopy, resulting from single-molecule analysis: A. side view of an MspA aggregate; B. top view of an MspA pore. For the original images, see [37].

4.3. The Advantages of MspA in Nanotechnology Compared to Other Proteins

Proteins are macromolecules with dimensions in the nanometer range and can be tailored to specific needs by site-directed mutagenesis. However, their use in nanotechnology is severely hampered because most proteins lose their structural integrity in a nonnative environment and, therefore, cannot be used in technical processes. The MspA porin from *M. smegmatis* is an extremely stable protein, which retains its channel structure even after boiling in 3% SDS or extraction with organic solvents [32]. This should considerably stabilize nanoparticles synthesized inside the channel. The same argument holds true for specific sensor functions inside the MspA channel. Longer template channels also allow the synthesis of longer nanowires extending their application potential. The hydrophobic surface of MspA allows its assembly into biomimetic membranes. Thus, MspA is the most stable channel-forming protein known thus far. Furthermore, MspA has an astonishing tendency to self-assemble into ordered structures on surfaces [37]. This permits the synthesis of nanoarray lattices as well as nanoparticles- and rods in a statistic distribution, embedded by a suitable hydrophobic surface.

Furthermore, MspA is the only mycobacterial porin that can be purified in milligram quantities. It is selectively extracted by boiling entire cells of *M. smegmatis* for 30 min in a buffer containing detergents and purified by chromatography to apparent homogeneity [86] (see Fig. 11).

4.4. Nanostructuring by Deposition of the MspA Porin on HOPG Surfaces

The generation of well-defined nanostructures by protein or macromolecule deposition on two-dimensional surfaces (area up to 1.0×10^{-4} m^2) represents a considerable advantage compared to state-of-the-art technologies. However, most conventional proteins lose their structural integrity in a nonnative environment and, therefore, cannot be used in technical processes. Due to the extreme stability of MspA, it can be used in the following simple and straightforward procedure for the nanopatterning of extended surface areas: MspA, dissolved in a buffer solution can be dispersed into droplets employing a simple sonication procedure. The formed droplets are allowed to reach a HOPG surface. Depending on the exact deposition conditions (temperature, MspA concentration, sonication duration and intensity, and the length of the curing procedure after deposition), various nanostructures were obtained at HOPG. Therefore, these experimental results offer a simple and straightforward approach to the nano-structuring of surfaces by the deposition of protein containing buffer droplets. At deposition and curing temperature of 30 °C, the generation of regular nanostructures, which feature nanochannels, was proven unambiguously by electron microscopy in combination with computer-assisted image analysis [39]. The formation of these highly desired nanochannels within the deposited layer proceeded most likely by the reconstitution of biologically active MspA-nanopores ($M \approx 80{,}000$ Da) within the codeposited MspA protein layer, which was formed by interaction of MspA monomers ($M \approx 20{,}000$ Da) [87].

Three different kinds of layer structures were created by depositing protein/buffer droplets on HOPG surfaces and independently analyzed by REM and TEM. The results are summarized and compared with each other in Figure 12. It becomes immediately clear that the temperature during this deposition of droplets and during the curing process (the surface was allowed to (nano)structure for at least 1 h of the chosen temperature) after deposition of protein/buffer mixtures determined the formed nanostructures at HOPG.

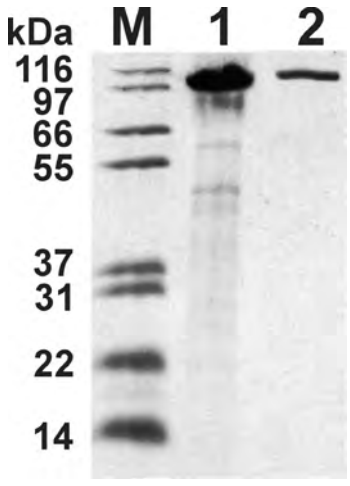


Figure 11. Protein samples were separated on a denaturing 10% polyacrylamide gel. Lanes: (M) molecular mass marker (200, 116.3, 97.4, 66.3, 55.4, 36.5, 31, 21.5, 14.4, 6 kDa); (1) 4 μg protein from MspA-containing fractions pooled after anion exchange chromatography. (2) 0.7 μg protein from MspA-containing fractions pooled after gel filtration chromatography. The proteins were stained with Commassie Blue.

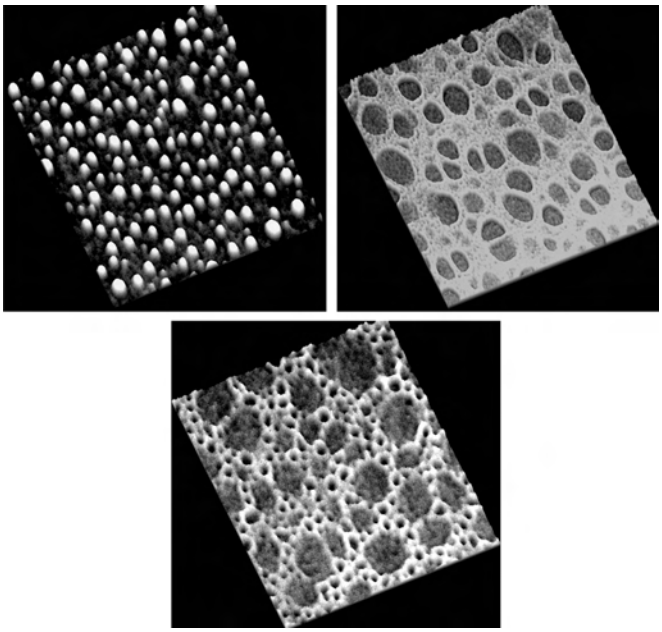


Figure 12. Three types of layer structures created by depositing protein/buffer droplets onto HOPG surfaces (surface plots created by IMAGE; further explanations are provided in the text.). The extension of each image is 110 nm × 110 nm.

After depositing MspA and curing at $T = 20$ °C, only isolated proteins and no channel structures were observed (Fig. 12, left image). The main diameter of the units in these images is 2.15 ± 0.40 nm. Deposition of the denatured MspA monomer resulted in similar structures (main diameter 2.14 ± 0.70 nm) indicating that MspA deposited mainly in the form of its monomer ($M = 20{,}000$ Da). Note that the denatured MspA monomer did not form any channels. This is in agreement with earlier findings, which demonstrated that the MspA-porin loses its ability to form nanochannels once it becomes the MspA monomer.

The raising of the deposition and the curing temperature to $T = 25$ °C, yielded in the formation of a supramolecular structure formed by MspA on the HOPG surface (Fig. 12, center image). The thickness of the protein layer was 4.5 ± 0.25 nm. Only indentations of the diameter of 9.2 ± 2.1 nm are found within the investigated structure. The residual thickness of the protein layer at the bottom of these dead-end channels was 1.4 ± 0.30 nm. This analysis revealed that no open channels were present in this supramolecular structure.

The dimensions of the MspA pore were determined by electron microscopy of negatively stained cell wall preparations of *M. smegmatis* and of purified MspA. The MspA pore has an inner and outer diameter of 2.5 and 10 nm, respectively, and a length of about 10 nm. Considering these dimensions of MspA, two interpretations of the supramolecular structure are possible: (i) The MspA pore might be intact, but is either completely perpendicular with respect to the HOPG surface or it is formed in an angle, which does not permit detection. (ii) The observed indentations might be formed by larger aggregates of interacting MspA monomers. Note that the TEM technique does not permit the elucidation of the exact structure of the filled bottom of these indentations. Only the integrated thickness of the carbon containing protein structure can be estimated.

After increasing the deposition and the curing temperature to $T = 30$ °C, real nanochannels possessing a regular diameter of 2.6 ± 0.50 nm are detectable as indicated by the appearance of the uncovered HOPG surface in the center region of the nanochannels (Fig. 12, right image). Considering the nonnative environment, the diameter of these channels is in excellent agreement with structure of MspA, demonstrating that the channel protein did not denature during the deposition procedure at $T = 30$ °C. It should be noted in this regard that the Tet repressor protein (TetR), which is a water-soluble DNA-binding protein and does not form channels [85], could not be deposited on a carbon surface using the sonication-deposition procedure described here. During sonication, a complete disruption of TetR was observed indicating that this nanostructuring method can only be applied to stable proteins. In Figure 13, a surface plot of the highly structured MspA layer on HOPG is shown using high-resolution TEM files as input. Geometric features down to a size of less than 2 nm are clearly discernible. MspA formed a layer of an average thickness of 5.8 ± 0.25 nm, which was attached to HOPG by means of hydrophobic interaction. Real nanochannels were formed by reconstitution of further MspA-monomers into this layer. Note that dead-end channels of a diameter of 7.2 ± 1.25 nm

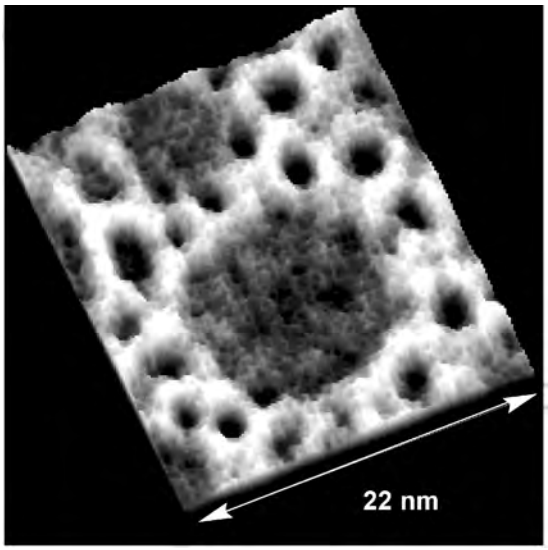


Figure 13. IMAGE surface plot a highly structured MspA layer on HOPG using high-resolution TEM as input.

were also detected within this highly complex layer structure. At the bottom of these indentations, the thickness of the MspA layer was reduced to 1.0 ± 0.2 nm.

In Figure 13, the TEM resolution is zoomed to its maximum limit. It can be clearly seen that the MspA nanochannels are open towards the HOPG surface at their lower ends. This permits an electrical contact between the interior of the porin channels, which are hydrophilic in nature, and the HOPG electrode. Due to the hydrophobic properties of the deposited MspA layer on HOPG, the protein structure proved to be stable for 24 h when immersed into H_2O.

The superposition of 50 nanopores is shown in Figure 14. A typical MspA pore featured a diameter of 3.1 ± 0.2 nm and a length of 9.5 ± 0.4 nm. The wall region surrounding the typical pore had an extension of 3.2 ± 0.4 nm. Whereas the average depth of the MspA layer was measured to be 5.8 ± 0.25 nm, the thickness of the wall region in close proximity to the MspA nanopores extended to 9.5 ± 1.0 nm.

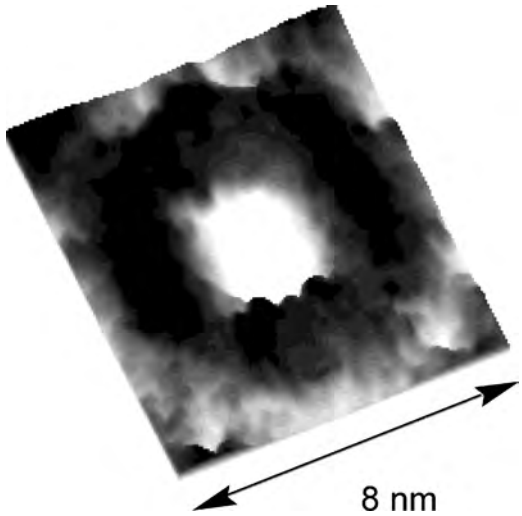


Figure 14. A typical MspA pore.

A more quantitative insight into the statistical size distribution of the nanochannels generated on HOPG was again provided by a detailed IMAGE analysis: For the calculation of Figure 15, the information from 10 different deposition experiments at $T = 30$ °C, was analyzed. As it is summarized in Figure 15, the statistic distribution of pore diameters deviates approximately by ± 0.50 nm from the mean value of 2.6 nm.

The generation of these highly ordered MspA nanostructures proceeded in a completely reproducible manner. In order to understand the remarkable process of MspA nanostructuring at hydrophobic surfaces, a look at the natural environment of this channel-forming protein might offer some additional information. The exceptionally thick (approximately 10 nm) and strongly hydrophobic membrane of *M. smegmatis* is the natural environment in which MspA nanochannels form. Based on the insight in the mycobacterial cell wall structure, the following interpretation of the observed channel-forming activity of MspA under technical conditions can be attempted: The biologically active form of the MspA is known to form nanochannels of a mass of approximately 80,000 Da. At $T = 20$ °C, the MspA monomer prevails at the HOPG surface indicating that no nanochannels survived the sonication procedure. The biologically inactive porin monomers, which are formed to various extents during the sonication process, possess hydrophobic as well as hydrophilic structure elements. Their amphiphilic nature permits various types of supramolecular interaction among originally identical protein monomers of the mass of 20,000 Da, as experienced by the formation of very different phases or layer structures, depending on the deposition and curing temperature.

At $T = 25$ °C, the MspA forms a highly complicated nanostructured layer at HOPG. It is still uncertain whether a fraction of the MspA nanochannels, which were present before sonication, are codeposited and therefore become a building element of this nanostructured layer. At $T = 30$ °C, a considerable fraction of the MspA nanochannels "survives" the sonication process most likely as a result from stronger hydrophobic interaction among the MspA subunits compared to lower temperatures. Presumably, the MspA oligomers reach the HOPG surface as nanochannels of an average diameter of 2.6 ± 0.50 nm and reconstitute in the supramolecular protein layer of codeposited MspA monomers at positions with the highest hydrophobicity. This hypothesis would explain the observation of enhanced thickness in the direct proximity of the MspA channel openings, because individual MspA pores reconstitute within an MspA protein layer of a different structure. Our future investigations will be aimed at a more detailed comprehension of the MspA deposition and channel-forming mechanisms in our biomimetic systems of high technical interest.

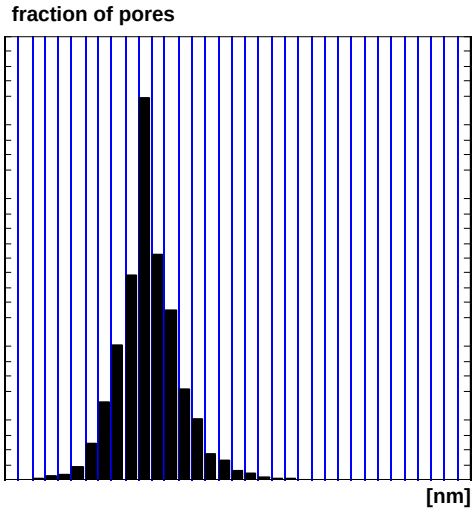


Figure 15. Quantitative analysis of the pore sizes formed by MspA on HOPG using IMAGE.

4.5. Nanostructuring of Carbon Surfaces By Deposition of a Channel-Forming Protein and Subsequent Polymerization of Methyl-Methacrylate-Prepolymers

As was previously pointed out, the MspA protein demonstrated its striking ability to form a wide range of nanostructures, including a nanochannel containing phase at a carbon surface (HOPG). However, the nanostructured MspA phases redissolved after several days in various aqueous buffers. This effect clearly hampers any electrochemical application using aqueous buffers, for instance, as a nanotechnological building kit for sensing applications. Therefore, we attempted to stabilize formed nanochannels using the hydrophobic homopolymer P(MMA). It could be expected that the phases formed using MspA and P(MMA) differ to various extents from the phases where MspA was used alone. Since very little is currently known about the prospect of self-structuring protein/polymer phases, the results are of great interest with respect to the generation of distinct shapes in the submicro and nanoscale. The experimental approach, as well as the IMAGE analysis procedures were very similar with respect to the studies on MspA on HOPG.

In Figure 16, the effect of the presence of MspA during the buffer deposition from the gas phase onto a carbon surface becomes clear. Without the presence of MspA, only disordered structures of P(MMA) (pre)polymers can be found, whereas distinct and reproducible (nano)structures

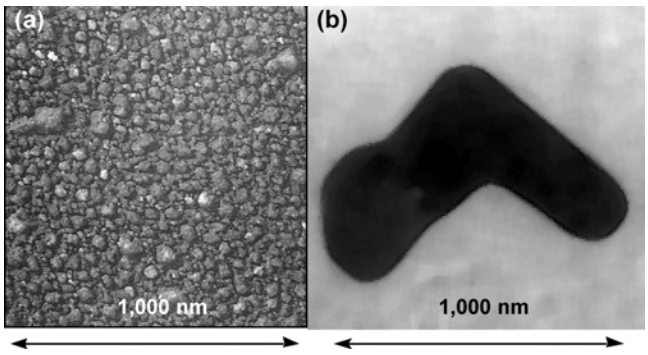


Figure 16. Effect of the deposition of P(MMA)/buffer (c = 2.25 μg/mL) (a) in comparison to MspA/P(MMA)/buffer (c(MspA) = 1.00 μg/mL, (c(PMMA/MMA) = 1.25 μg/mL)); (b) at carbon surfaces (TEM image).

are formed, if MspA is present. This finding provides the first evidence for the importance of MspA for the generation of ordered structures at surfaces.

In order to reach the goal of generating defined (nano)structures at a carbon surface, the general deposition procedure of a MspA/P(MMA) buffer was systematically varied to explore its possibilities [39]. Note that this procedure using buffer sonication and the following deposition of microdroplets at the carbon surface differs from a conventional adsorption procedure, because no equilibrium between MspA/P(MMA) and the surfaces exists. Therefore, the exposition time of the surface toward the MspA/P(MMA) buffer is the decisive factor for the amount of MspA/P(MMA) deposited. Again, the deposition and curing temperature proved to be the most decisive factor. And again, *remarkable and reproducible nanopatterns* at HOPG were discovered.

4.6. Formation of Microletters

When a depositon time of 125 ± 5 s and a deposition temperature of 30 ± 2 °C is chosen, the formation of "microletters" is observed. Typical "letters" are shown in Figure 17. This particular occurrence (thickness > 100 nm) is typical for a deposited pattern that was only investigated by TEM for a short time (less than 60 s).

4.7. Formation of Dendrites and Stars

In Figure 18, dendrites and "stars" generated by deposition of MspA for 175 ± 5 s at 40 ± 5 °C at the chosen carbon surface are shown. The comparison of the occurrence of this micropattern at very low TEM exposition times ($t < 60$ s) and after 300 s of electron beam treatment reveals polymerization of MMA included in the formed micropattern. The average thickness of the formed dendrites and stars was approximately 110 to 120 nm in the beginning of the TEM characterization, whereas the thickness decreased to 25–30 nm after 300 s of continuous electron beam exposure. In addition to the observed polymerization process, evaporation of water and MMA molecules into the high vacuum can be detected. Consequently, the diameter of the TEM beam of approximately 3000 nm can be seen in Figure 18 (300 s). However, this behavior is typical for all micropatterns investigated in this study. In all cases, the geometries initially formed during MspA deposition are retained during the electron beam treatment. This finding makes this simple deposition process of particular interest for the lateral nanostructuring of surfaces.

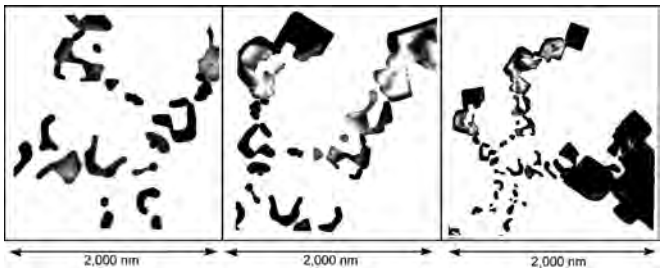


Figure 17. Microletters formed by deposition of MspA/P(MMA)/buffer at the carbon surfaces (TEM image, three different deposition experiments).

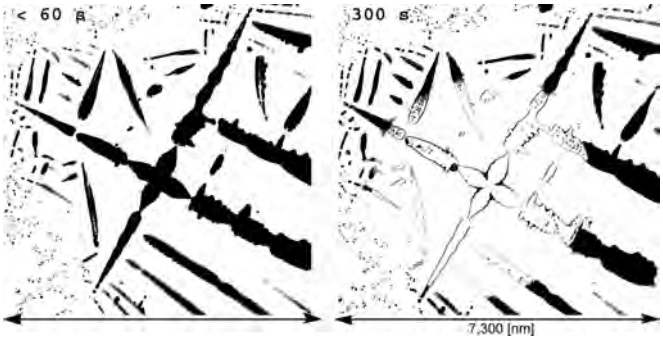


Figure 18. Dendrites and stars formed by deposition of MspA/P(MMA)/buffer at the carbon surface (TEM image) after<60 s and 300 s exposure to the electron beam.

4.8. Electron Beam-Induced Polymerization of MMA in Microletters and Quantitative Analysis of the Micropattern by Using IMAGE

Figure 19 presents a series of images of a microletter, which was continuously treated by the TEM electron beam. Whereas the letter appears to be relatively homogeneous

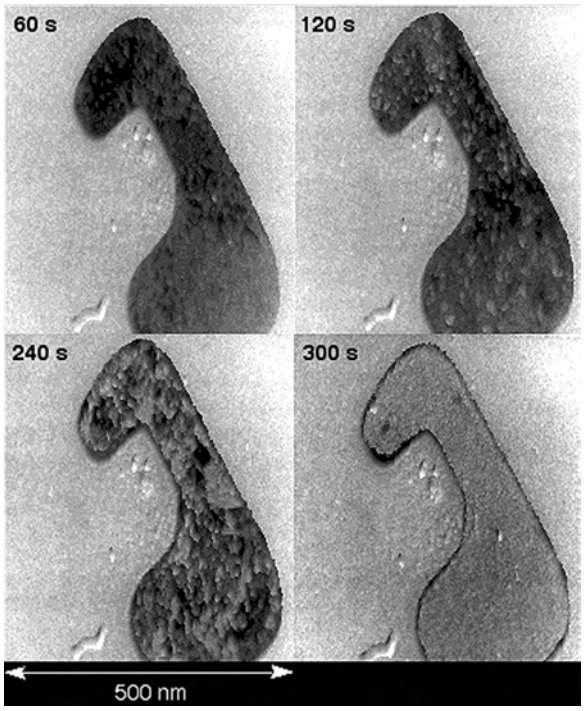


Figure 19. Surface plots of a typical microletter formed by deposition of MspA/P(MMA)/buffer at the carbon surface, calculated by IMAGE using the TEM images as input.

at the beginning of the electron beam fixation process, a patchwork of black and gray spots is present after 120 s and 240 s. However, after 300 s the interior of the microletter becomes homogeneously gray, whereas a sharp edge remains, which represents the exact geometric shape of the letter immediately after its deposition on the carbon surface. The analysis of the TEM images, using the program package IMAGE employing the calibration parameters obtained by the deposition, reveals striking insights into this fixation process. In the beginning of the electron beam exposure, the inner region of the microletter structure is smooth and possesses a thickness of 100–120 nm. In the course of the progressing reaction, holes with a diameter of 2–5 nm appear randomly distributed throughout the whole interior of the microletters. On the bottom of these holes exists an organic layer of a thickness of approximately 25 nm. In conclusion, the depth of the formed holes is approximately 75–95 nm. These hole regions grow consequently until most of the interior consists of an organic layer of 20–25 nm thickness. The surface of this layer structure is obviously rough.

A strong possibility exists that channel-like structures possessing a diameter of approximately 1–2 nm are randomly oriented within this layer. From our IMAGE analysis, it appears that holes, cavities, and/or channels exist at particular locations, where organic material is missing. However, using TEM, we could not elucidate whether channels exist, which are open to the surface of the microstructure. According to the IMAGE calibration procedure, the mass deposited in the interior of the microletters is approximately 2.7×10^{-9} g/cm^2.

In contrast to the interior of the microletters, their border regions form a wall-type structure, which is not affected by the continuous electron beam exposure. The diameter of these walls is approximately 65 nm at the bottom. The deposited mass at the wall regions is considerably higher (2.1×10^{-8} g/cm^2) than in the interior. After 300 s, only minor changes occurred in the fixed microletters at the carbon surface.

The wall regions that form the microstructure's boundaries might be formed from MspA and PMMA, which form hydrophobic aggregates. The chemical composition of these aggregates differs depending on the deposition conditions, such as MspA and PMMA—concentrations and temperature. Neither the deposition of MspA nor PMMA prepolymers alone at the reported experimental conditions resulted in the formation of the microstructures reported here. Since PMMA is highly polydispers ($P_d = 5.2$), it cannot be responsible for the patterning process. Therefore, it is very likely that both MspA and PMMA are the main building components of the wall-type structures: hydrophobic aggregates of MspA and PMMA develop during the deposition of the MspA/P(MMA) buffer microdroplets on the carbon surface. The hydrophobicity of the carbon surface itself may support the formation of these hydrophobic aggregates. Within the formed microstructures, hydrophilic cavities containing mostly MspA in aqueous buffer and a hydrophobic microphase consisting mostly of MMA and polymerization initiator coexist. Perkovic and McConnell succeeded in the mathematical modeling of complex two-dimensional patterns in mixed surfactant bilayers, and also in protein layers, by using a line tension—dipole repulsions model [88]. The observed equilibrium shape of a constant area monolayer was explained from the competition between the line tension of the domain boundary and the dipole-dipole repulsions between the molecular dipoles. Unfortunately, our much more complex and three-dimensional nanostructures cannot be explained as of yet by applying this model. However, we envision the mechanisms that lead to the formation of the distinct pattern during deposition to be very similar to those developed by McConnell in the two-dimensional phase.

Upon electron beam exposure and consequent heating of the microstructures, water as well as some of the MMA monomer evaporates in the high vacuum. The electron beam impact, as well as the increased temperature caused by TEM, leads to either anionic or radical MMA polymerization [89]. The gel effect is well known during MMA polymerization—at enhanced temperature the reaction rate of PMMA formation increases rapidly until a gel is formed from PMMA and the nonreacted MMA. The formation of the gel results in a drastic *decrease* of the polymer volume.

Both the occurrence of the gel effect and the evaporation of H_2O or MMA at the same time may explain the appearance of "holes" during electron beam impact. This process continues until neither H_2O nor MMA remains in the interior of the microstructures. Upon further observation, it was evident that the wall regions do not exhibit the "shrinking effect" during electron beam impact; therefore, we can conclude that only very little "free" H_2O and/or MMA is present in these regions after deposition.

5. CONCLUSION

S-layer proteins and the channel porins from Gram-positive bacteria represent two complementary approaches to new materials, processes, and device architectures for information recording and systems. This approach consists of the formation of protein-based layers (S-layer lattices or hydrophobic (bi)layers, in which MspA reconstitutes) on semiconducting and conducting surfaces. Other surface types, such as liposomes, vesicles, or bacterial membranes will be used in future applications, which can be regarded as spin-offs of the nanotechnology endeavor. Whereas the enormous potential of the S-layer proteins is known to an increasing scientific audience, the channel porin from *Mycobacterium smegmatis* (MspA), which represents a strong challenge to carbon nanotubes as host material for nanodots and nanowires, will be increasingly employed as a nanotechnological building kit.

GLOSSARY

Absorption Movement of ions and water into as organism as a result of metabolic processes, frequently against an electrochemical potential gradient (active) or as a result of diffusion along an activity gradient (passive).

N-acetylglucosamine and N-acetylmuramic acid Sugar derivatives in the peptidoglycan layer of bacterial cell walls.

Adsorption Process by which atoms, molecules, or ions are taken up and retained on the surfaces of solids by chemical or physical binding.

Aliphatic Organic compound in which the main carbon structure is a straight chain.

Amino group An -NH_2 group attached to a carbon skeleton as in the amines and amino acids.

Archaea Evolutionarily distinct group (domain) of prokaryotes consisting of the methanogens, most extreme halophiles and hyperthermophiles, and thermoplasma.

Archaebacteria Older term for the Archaea.

Aromatic Organic compounds which contain a benzene ring, or a ring with similar chemical characteristics.

Bacillus Bacterium with an elongated, rod shape.

Bacteria All prokaryotes that are not members of the domain Archaea.

Biosensor A device designed to detect a specific biological molecule, system of biological molecules, or biologically produced signal.

Biotechnology Use of living organisms to carry out defined physiochemical processes having industrial or other practical application.

Cell Fundamental unit of living matter.

Cell membrane See cytoplasmic membrane.

Cell wall Layer or structure that lies outside the cytoplasmic membrane; it supports and protects the membrane and gives the cell shape.

Covalent Nonionic chemical bond formed by a sharing of electrons between two atoms.

Cyanobacterium Prokaryotic, oxygenic phototrophic bacterium containing chlorophyll a and phycobilins, formerly the "blue-green algae."

Cytoplasm Cellular contents inside the cell membrane, excluding the nucleus.

Cytoplasmic membrane Selectively permeable membrane surrounding the cell's cytoplasm.

Deoxyribonucleic acid (DNA) A naturally occurring polymer composed of deoxyribonucleotides. Although commonly found as a double-stranded, right-handed helix it can adopt one-, three-, and four-stranded forms having a variety.

Electrophoresis Separation of charged molecules, such as nucleic acids, in an electrical field.

Eubacteria Old term for the bacteria.

Eukarya Phylogenetic Domain containing all eukaryotic organisms.

Fluorescent Able to emit light of a certain wavelength when activated by light of a shorter wavelength.

Fluorescent antibody Antiserum conjugated with a fluorescent dye, such as fluorescein or rhodamine.

Gel Inert polymer, usually made of agarose or polyacrylamide, that separates macromolecules such as nucleic acids or proteins during electrophoresis.

Gram stain Differential stain that divides bacteria into two groups, Gram-positive and Gram-negative, based on the ability to retain crystal violet when decolorized with an organic solvent such as ethanol. The cell wall of Gram-positive bacteria consists chiefly of peptidoglycan and lacks the outer membrane of Gram-negative cells.

Hydrogen bond A weak polar bond formed between two electron-rich atoms, one of which is covalently bonded to a hydrogen atom.

Immobilization Conversion of an element from the inorganic to the organic form in microbial or plant biomass.

Immunofluorescence Technique to visualize specific antibodies and any attached homologous antigens by means of conjugating the antibodies to a fluorescent dye.

Ions Atoms, groups of atoms, or compounds, that are electrically charged as a result of the loss of electrons (cations) or the gain of electrons (anions).

Luminescence Generation of light.

Macromolecule Large molecule formed from the connection of a number of small molecules.

Microbiology Study of microorganisms.

Molecule Result of two or more atoms combining by chemical bonding.

Molecular modeling A computer-aided design process in which computational chemistry software programs are used to create models of macromolecules based on electronic structure, and molecular mechanics calculations coupled with 3D experimental data.

Mutant Organism, population, gene, or chromosome that differs from the corresponding wild type by one or more base pairs.

Mutation A change in the information storage mechanism of an evolving system that alters the information it contains but may or not alter the form or function of the system.

Nanopore Pore having dimensions measured in nanometers.

Peptidoglycan Rigid layer of cell walls of bacteria, a thin sheet composed of N-acetylglucosamine, N-acetylmuramic acid, and a few amino acids. Also called murein.

pH Negative logarithm of the hydrogen ion activity. The degree of acidity (or alkalinity) of a soil as determined by means of a glass or other suitable electrode or indicator at a specified moisture content or soil-water ratio, and expressed in terms of the pH scale.

Polar Possessing hydrophilic characteristics and generally water soluble.

Porin A protein which forms at least one channel in a bacterial membrane.

S-layer proteins Proteins at the outside of many bacterial membranes.

REFERENCES

1. C. L. Claeys and H. E. Maes, "Microelectronics, Microsystems and Nanotechnology," Papers presented at MMN 2000 Nov. 20–22, 91–102 (2001).
2. M. Mitsuishi, T. Li, and T. Miyashita, *Mol. Crys. Liq. Crys. Science Tech.* 2001, 269 (2001).
3. Y. Seo, K. Lee, M. Yi, E. Seo, B. K. Choi, O. Kim, I. Raptis, P. Argitis, and M. Hatzakis, *Microelectron. Eng.* 46, 461 (1999).
4. W. H. Bruenger, R. Kaesmaier, H. Loeschner, and R. Springer, "Materials Research Society Symposium Proceedings," 2001, pp. 1–12.

5. V. F. Puntes, Z. Konya, C. Erdonmez, J. Zhu, G. A. Somorjao, and A. P. Alivisatos, "Materials Research Society Symposium Proceedings," 2002, pp. 241–246.
6. V. K. Varadan, "Proceedings of SPIE-The International Society for Optical Engineering," 2001, pp. 25–45.
7. S. Link and M. A. El-Sayed, *J. Phys. Chem. B.* 103, 8410 (1999).
8. M. T. Reetz, W. Helbig, S. A. Quaiser, U. Stimming, N. Breuer, and R. Vogel, *Science* 267, 367 (1995).
9. S. Luscher, A. Fuhrer, R. Held, T. Heinzel, K. Ensslin, M. Bichler, and W. Wegscheider, *Microelectron. J.* 33, 319 (2002).
10. R. T. Pötzschke, G. Staikov, W. J. Lorenz, and W. Wiesbeck, *J. Electrochem. Soc.* 146, 141 (1999).
11. U. B. Sleytr, M. Sara, D. Pum, and B. Schuster, *Nano-Surface Chem.* 333 (2002).
12. U. B. Sleytr, M. Sara, and D. Pum, *Supramol. Polymers* 177 (2000).
13. D. Pum, M. Sara, and U. B. Sleytr, *J. Vac. Sci. Technol. B.* 7, 1391 (1989).
14. M. Tinkham, D. Davidovic, D. C. Ralph, and C. T. Black, *J. Low Temp. Phys.* 118, 271 (2000).
15. L. D. Burke and P. F. Nugent, *Gold Bull.* 31, 39 (1998).
16. G. Y. Tseng and J. C. Ellenbogen, *Science* 294, 1293 (2001).
17. D. Srivastava, M. Menon, and K. Cho, *Compt. Sci. Eng.* 3, 42 (2001).
18. S. H. Bossmann, C. Turro, C. Schnabel, M. R. Pokhrel, L. M. J. Payawan, and M. Wörner, *J. Phys. Chem. B.* 105, 5374 (2001).
19. C. A. J. Foss, *Metal Nanoparticles* 119 (2002).
20. A. Huczko, *Appl. Phys. A*. 70, 365 (2000).
21. Z. Konya, V. F. Puntes, I. Kiricsi, J. Zhu, P. Alivisatos, and G. A. Somorjai, *Catal. Lett.* 81, 137 (2002).
22. M. Sara and U. B. Sleytr, *Prog. Biophys. Molec. Biol.* 65, 83 (1996).
23. H. Nikaido, S. H. Kim, and E. Y. Rosenberg, *Mol. Microbiol.* 8, 1025 (1993).
24. G. E. Schulz, *Curr. Opin. Struct. Biol.* 6, 485 (1996).
25. T. Schirmer, *J. Struct. Biol.* 121, 101 (1998).
26. R. Koebnik, K. P. Locher, and P. V. Gelder, *Mol. Microbiol.* 37, 239 (2000).
27. T. Lichtinger, A. Burkovski, M. Niederweis, R. Krämer, and R. Benz, *Biochemistry* 37, 15024 (1998).
28. F. G. Riess, T. Lichtinger, R. Cseh, A. F. Yassin, K. P. Schaal, and R. Benz, *Mol. Microbiol.* 29, 139 (1998).
29. E. Maier, G. Polleichtner, B. Boeck, R. Schinzel, and R. Benz, *J. Bacteriol.* 183, 800 (2001).
30. X. Jiang, M. A. Payne, Z. Cao, S. B. Foster, J. B. Feix, S. M. C. Newton, and P. E. Klebba, *Science* 276, 1261 (1997).
31. W. Welte, K. Diederichs, M. Przybylski, M. O. Glocker, R. Benz, and J. Breed, *NATO ASI Series C Mathematical and Physical Sciences—Advanced Study Institute* 510, 239 (1998).
32. A. Mathes and H. Engelhardt, *Biophys. J.* 75, 1255 (1998).
33. D. J. Müller and A. Engel, *J. Mol. Biol.* 285, 1347 (1999).
34. T. Schirmer, T. A. Keller, Y.-F. Wang, and J. P. Rosenbusch, *Science* 267, 512 (1995).
35. W. Im and B. Roux, *J. Mol. Biol.* 319, 1177 (2002).
36. M. Niederweis, S. Ehrt, C. Heinz, U. Klöcker, S. Karosi, K. M. Swiderek, L. Riley, and R. Benz, *Mol. Microbiol.* 33, 933 (1999).
37. H. Engelhardt, C. Heinz, and M. Niederweis, *J. Biol. Chem.* 227, 37567 (2002).
38. C. Weiss-Wichert, M. Smetazko, M. Valina-Saba, and T. Schalkhammer, *J. Biomol. Screen.* 2, 11 (1997).
39. M. Niederweis, C. Heinz, K. Janik, and S. H. Bossmann, *Nano Letters*, 1, 169 (2001).
40. U. B. Sleytr, M. Sara, and D. Pum, *Microelectron. Eng.* 9, 13 (1989).
41. H. Bahl, H. Scholz, N. Bayan, M. Chami, G. Leblon, T. Guli-Krzywicki, E. Shechter, A. Fouet, S. Mesnage, E. Tosi-Couture, P. Gounon, M. Mock, E. C. D. Macario, A. J. L. Macario, L. A. Fernandez-Herrero, G. Olabarria, J. Berenguer, M. J. Blaser, W. Lubitz, B. Kuen, M. Sara, P. H. Pouwels, C. P. A. M. Kolen, H. J. Boot, A. Palva, M. Truppe, S. Howorka, G. Schroll, S. Lechleitner, and S. Resch, *FEMS Microbiology Rev.* 20, 47 (1997).
42. B. Schuster and U. B. Sleytr, *Rev. Mol. Biotech.* 74, 234 (2000).
43. M. Sara and U. B. Sleytr, *J. Bacteriol.* 182, 859 (2000).
44. B. Kuen and W. Lubitz, *Crystalline Bacterial Cell Surface Proteins* 77 (1996).
45. M. Sara and E. M. Egelseer, *Crystalline Bact. Cell Surf. Proteins* 103 (1996).
46. P. Messner, *Crystalline Bact. Cell Surf. Proteins* 35 (1996).
47. U. B. Sleytr, H. Bayley, M. Sara, A. Breitwieser, S. Kuepcue, C. Mader, S. Weigert, F. M. Unger, P. Messner, B. Jahn-Schmid, B. Schuster, D. Pum, K. Douglas, N. A. Clark, J. T. Moore, T. A. Winningham, S. Levy, I. Frithsen, J. Pankov, P. Beale, H. P. Gillis, D. Choutov, and K. P. Martin, *FEMS Microbiol. Rev.* 20, 151 (1997).
48. S. Kupcue, M. Sara, S. Weigert, and U. B. Sleytr, *Adv. Filtration Sep. Technol.* 7, 416 (1993).
49. B. Schuster, D. Pum, M. Sara, O. Braha, H. Byley, and U. B. Sleytr, *Langmuir* 17, 499 (2001).
50. S. Kupcue, M. Sara, and U. B. Sleytr, *J. Membrane Sci.* 61, 167 (1991).
51. S. Howorka, M. Sara, Y. Wang, B. Kuen, U. B. Sleytr, W. Lubitz, and H. Byley, *J. Biol. Chem.* 275, 37876 (2000).
52. C. Mader, S. Kupcue, U. B. Sleytr, and M. Sara, *Biochimica et Biophysica Acta* 1463, 142 (2000).
53. S. Kupcue, C. Mader, and M. Sara, *Biotech. Appl. Biotech.* 21, 275 (1995).
54. B. Schuster and U. B. Sleytr, *Rev. Mol. Biotech.* 74, 234 (2000).
55. D. Pum, A. Neubauer, U. B. Sleytr, S. Pentzien, S. Reetz, and W. Kautek, *Berichte der Bunsen-Gesellschaft* 101, 1686 (1997).
56. A. Neubauer, S. Pentzien, S. Reetz, W. Kautek, D. Pum, and U. B. Sleytr, *Sensors and Actuators, B* B40, 231 (1997).
57. D. Pum, A. Neubauer, E. Gyorvary, M. Sara, and U. B. Sleytr, *Nanotechnology* 11, 100 (2000).
58. D. Pum, G. Stangl, C. Sponer, K. Riedling, P. Hudek, W. Fallmann, and U. B. Sleytr., *Microelectron. Eng.* 35, 297 (1997).
59. N. J. Turro, "Modern Molecular Photochemistry 1," University Science Books, Mill Valley, CA, (1991).
60. W. Shenton, D. Pum, U. B. Sleytr, and S. Mann, *Nature* 389, 585 (1997).
61. S. Dieluweit, D. Pum, and U. B. Sleytr, *Supramol. Sci.* 5, 15 (1998).
62. C. Stathopoulos, *Membr. Cell. Biol.* 13, 3 (1999).
63. M. Nasseau, Y. Boublik, W. Meier, M. Winterhalter, and D. Fournier, *Biotechnol. Bioeng.* 75, 615 (2001).
64. X. Jiang, M. A. Payne, Z. Cao, S. B. Foster, J. B. Feix, S. M. C. Newton, and P. E. Klebba, *Science* 276, 1261 (1997).
65. F. G. Riess, T. Lichtinger, R. Cseh, A. F. Yassin, K. P. Schaal, and R. Benz, *Mol. Microbiol.* 29, 139 (1998).
66. W. Welte, K. Diederichs, M. Przybylski, M. O. Glocker, R. Benz, and J. Breed, *NATO ASI Series C Mathematical and Physical Sciences—Advanced Study Institute* 510, 239 (1998).
67. F. P. Thinnes, K. P. Hellmann, T. Hellmann, R. Merker, C. Schwarzer, G. Walter, H. Götz, and N. Hilschmann, *Mol. Genet. Metab.* 69, 240 (2000).
68. A. Mathes and H. Engelhardt, *Biophys. J.* 75, 1255 (1998).
69. Z. Markovic-Housley and R. M. Garavito, *Biochim. Biophys. Acta* 869, 158 (1986).
70. P. S. Phale, A. Philippsen, C. Widmer, V. P. Phale, J. P. Rosenbusch, and T. Schirmer, *Biochemistry* 40, 6319 (2001).
71. W. Meier, C. Nardin, and M. Winterhalter, *Angew. Chem., Int. Ed.* 39, 4599 (2000).
72. C. Nardin and W. Meier, *Rev. Mol. Biotech.* 90, 17 (2002).
73. A. Graff, M. Winterhalter, and W. Meier, *Langmuir* 17, 919 (2001).
74. M. Nasseau, Y. Boublik, W. Meier, M. Winterhalter, and D. Fournier, *Biotechnol. Bioeng.* 75, 615 (2001).
75. S. W. Cowan, T. Schirmer, G. Rummel, M. Steiert, R. Ghosh, R. A. Pauptit, J. N. Jansonius, and J. P. Rosenbusch, *Nature* 258, 727 (1992).

76. C. Nardin, S. Thoeni, J. Widmer, M. Winterhalter, and W. Meier, *Chem. Commun.* 1433 (2000).
77. P. J. Brennan and H. Nikaido, *Annu. Rev. Biochem.* 64, 29 (1995).
78. H. Nikaido, S. H. Kim, and E. Y. Rosenberg, *Mol. Microbiol.* 8, 1025 (1993).
79. J. Liu, E. Y. Rosenberg, and H. Nikaido, *Proc. Natl. Acad. Sci.* 92, 11254 (1995).
80. J. Liu, C. E. Barry, G. S. Besra, and H. Nikaido, *J. Biol. Chem.* 271, 29545 (1996).
81. B. A. Dmitriev, S. Ehlers, E. T. Rietschel, and P. J. Brennan, *Int. J. Med. Microbiol.* 290, 251 (2000).
82. J. Trias, V. Jarlier, and R. Benz, *Science* 258, 1479 (1992).
83. J. Trias and R. Benz, *Molec. Microbiol.* 14, 283 (1994).
84. C. Stahl, S. Kubetzko, I. Kaps, S. Seeber, H. Engelhardt, and M. Niederweis, *Molec. Microbiol.* 40, 451 (2001).
85. D. Schnappinger and W. Hillen, *Arch. Microbiol.* 165, 359 (1996).
86. C. Heinz and M. Niederweis, *M. Anal. Biochem.* 285, 113 (2000).
87. M. Niederweis, C. Heinz, K. Janik, and S. H. Bossmann, Nanoletters, 2, 1263 (2002).
88. S. Perkovic and H. M. McConnell, *J. Phys. Chem. B.* 101, 381 (1997).
89. S. H. Bossmann, S. Teixeira, and A. M. Braun, "MMA-Photopolymerisation 1," Springer-Verlag, Berlin, 2000.

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanotribology

T. Coffey, J. Krim

North Carolina State University, Raleigh, North Carolina, USA

CONTENTS

1. INTRODUCTION

Tribology (from the Greek word "tribos," which means "to rub") is the study of friction, lubrication, and wear, an inherently challenging field requiring much interdisciplinary cooperation. By most recent estimates, improved attention to tribology would save developed countries up to 1.6% of their gross national product, or over $100 billion annually in the United States alone [1]. The magnitude of the financial loss associated with friction and wear arises from the fact that entire mechanical systems, be they coffee makers or automobiles, are frequently scrapped when only a few of their parts are badly worn. In the case of an automobile, the cost of the energy consumed in its manufacture is equivalent to that required to operate it for 100,000 mi [2].

Macroscopic tribological behavior is known to be highly sensitive to the details of interfacial contact at the atomic scale. For example, it is well known that adsorption of molecularly thin coatings on contacting surfaces can reduce their macroscopic friction coefficients by orders of magnitude. Investigations of the origins of such macroscopic tribological behavior should ideally be performed at both atomic length scales (10^{-9} m) and time scales (10^{-12}–10^{-9} s). Such studies, however, have proven to be extremely challenging, as they are performed on buried interfaces that are extremely difficult to experimentally probe.

With the advent of new experimental techniques emerging in the 1980s, a new field involving friction studies at atomistic length and time scales, nanotribology, emerged. The increased interest in nanotribology sparked a variety of discussions and debates concerning the nature of the atomic-scale mechanisms that dominate the dissipative process by which mechanical energy is transformed into heat. Decades-old suggestions that friction could originate in sliding-induced excitation of atomic lattice vibrations (phonons, i.e., sound waves) were reexamined, and suggestions for electronic contributions (associated with electronic motion excited by the sliding process) to friction also emerged. The renewed interest in nanotribology is now beginning to impact mainstream macrotribological considerations [3, 4]. Meanwhile, the results are most relevant to applications such as solid–liquid interfaces, where the complicating factors associated with asperity contacts are less of an issue, and to the MEMS community [5], where machine components with astoundingly small dimensions are rapidly approaching the length scales routinely probed by the nanotribological community. In the wider arena of nanotechnology, energy dissipation and conversion issues are integral to successful device design, construction, and/or operation [6]. This is obviously true when considering device energy efficiency from a profitability point of view, but equally compelling is the fact that any heat released in an inefficient transport process is available to destroy device function: if any portion melts, the device itself is destroyed.

While advances in recent years have been significant, much work lies ahead. This chapter provides an introduction to some of the important topics and concepts in nanotribology, with a focus on friction at the atomic scale. Key experiments and discoveries are detailed, including the experimental techniques employed for studies of nanotribology, the fundamental origins of friction, phononic and electronic contributions to friction, static friction, frictional anisotropy, and the velocity dependence of friction at atomic length and time scales.

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (869–877)

2. THE FUNDAMENTAL ORIGINS OF MACROSCOPIC FRICTION

When an object moves along a surface or through a viscous liquid or gas, the forces resisting its motion are referred to as friction. Frictional forces are nonconservative, converting the kinetic energy of materials in sliding contact to internal energy. Thus, if an object is given an initial velocity along a horizontal surface, the temperature of both the object and the surface upon which it slides will increase as the friction resisting its motion brings it to a stop. If an object is initially at rest, a minimum force must be applied to overcome "static friction," giving rise to the familiar phenomenon of an object "jumping ahead" at the instant that sliding is initiated. Static friction, in contrast to kinetic friction, is associated with neither energy "loss" nor sample heating. It is entirely absent for bodies moving through viscous fluids, "viscous friction" being parameterized at low velocities by

$$F \propto v \tag{1}$$

where v is the velocity of the object through the fluid.

For objects sliding along solid surfaces, friction is generally well described on the macroscopic scale by Amontons' law [7]:

$$F = \mu N \tag{2}$$

as reported in most introductory physics textbooks. The "normal load" N in Eq. (2) is the force that presses surfaces in sliding contact together, and μ is the "coefficient of friction." Amontons also reported that the friction force was independent of the apparent area of contact: a small block experiences as much friction as a large block of the same material as long as their weights are equal. A third law, attributed to the French physicist Charles Augustin Coulomb, is frequently included with those of Amontons: the friction force is independent of velocity for ordinary sliding speeds.

One common misconception about how friction originates is that it is always associated with interfacial roughness on some suitably magnified length scale. Surface roughness has, in fact, been ruled out for decades as a viable mechanism for everyday friction. Automobile makers, for example, routinely find that the friction between two surfaces is sometimes less if one of the surfaces is rougher than the other [8]. Furthermore, friction can increase if two surfaces are made smoother. In cold welding, for example, highly polished metals stick together firmly.

Interfacial roughness is, in fact, closely linked to Amontons' law in the sense that atomically flat surfaces in sliding contact are not described by it. The vast majority of surfaces are, of course, not atomically flat. When two such "rough" surfaces touch, contact between them takes place only at their asperities. The actual microscopic area of contact A_c is thus much less than the apparent area of contact A, perhaps by a factor of 10^4. When the surfaces are forced to slide over each other, new contact regions are continuously formed, while others are severed. If the true area of contact is constant on average, and the contacting junctions all have the same shear strength s, then the friction force is

$$F = A_c s \tag{3}$$

Equation (3) has been experimentally verified by a broad range of experimental geometries and a wide range of experimental techniques [9–13]. Enachescu et al. [13], for example, measured the friction force of a tungsten–carbide probe sliding on a boron-doped diamond surface. The friction force for this system was then compared to the local contact conductance, which revealed the true area of contact, and was observed to be linearly proportional to the true area of contact.

The relations $F = A_c s$ and $F = \mu N$ can be linked if one succeeds at establishing the dependence between the true contact area A_c and the normal load N. Many experiments have been performed with this goal in mind, but direct measures of true contact area are extremely difficult to achieve. Theoretical models of contact have therefore been routinely employed to aid in the interpretation of indirect experimental measures. By using contact mechanics models between two spheres, such as those derived by Johnson, Kendall, and Roberts (JKR) [14], Derjaguin, Muller, and Toporov (DMT) [15], and Hertz [16], it has been shown that $A_c \sim N^{2/3}$ for perfect elastic deformation and $A_c \sim N^1$ for plastic deformation. Experimental investigations of contacting spheres have found that, for N^n, n is closer to 2/3 for rubber, wood, plastic, and textiles, and n is closer to 1 for more brittle materials such as glass, diamond, and rock salt. And although only a limited number of materials exhibit $A_c \sim N^1$ for two spheres in contact, a far wider range of materials displays this dependence when multiasperity contact is considered [2]. Note that the true area of contact technically does not drop to zero for zero external load, as molecular adhesive forces are still present.

The dependence of friction on true contact area describes the macroscopic system's response to changes in load, but sheds little light on the physical origins of friction. This information is contained in the shear stress s, the frictional force per unit area of known contact. Bowden and Tabor's group explored this issue extensively, and were able to rule out the possibility that strong bonding at asperity contacts followed by wear was the physical mechanism underlying friction [17]. Indeed, Tabor and Israelachvili, working with an adapted version of the surface forces apparatus (SFA), were the first to provide conclusive experimental evidence that friction could occur in the *total absence* of wear [18, 19].

3. EXPERIMENTAL TECHNIQUES FOR STUDYING NANOTRIBOLOGY

There are three main experimental techniques for studying nanotribology. They are atomic force microscope or lateral force microscope (AFM or LFM), the surface forces apparatus (SFA), and the quartz crystal microbalance (QCM). These three techniques cover a wide range of length scales, time scales, and shear stress (see Fig. 1). LFM is characterized by atomic-scale length measurements, slow sliding speeds (on the order of 1 μm/s), or long time scales, and high shear stress (from 10^4 to 10^{10} N/m^2). SFA has comparatively long length and time scales and medium-range shear stresses (from 10^3 to 10^7 N/m^2). QCM has atomic length scales, short time scales (slip times on the order of nanoseconds), and low shear stress (from 10^{-2} to 10^2 N/m^2).

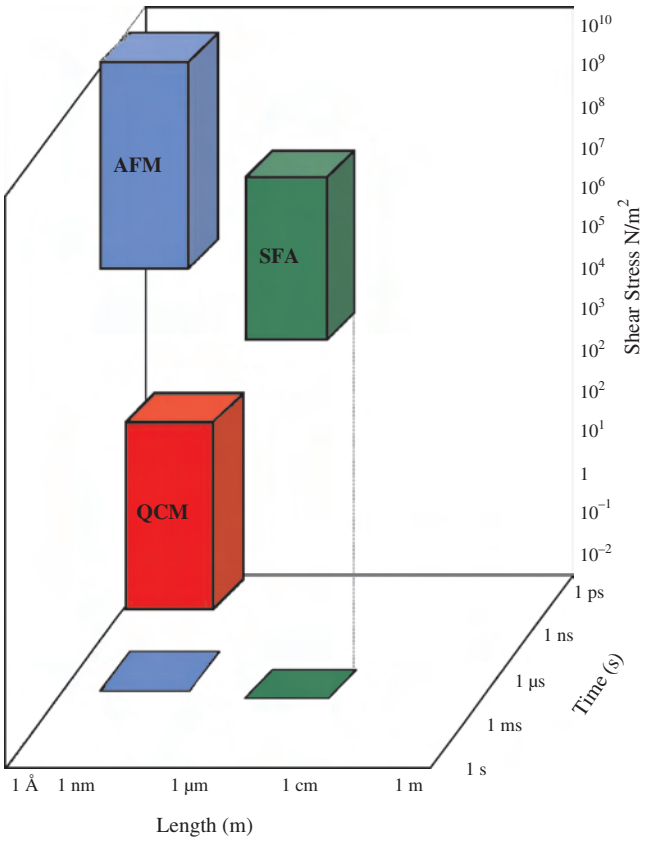


Figure 1. Length scales, time scales, and shear stresses of present-day nanotribological techniques.

The AFM was adapted for atomic-scale friction measurements in 1987 by Mate et al. [20]. An AFM typically consists of a sharp tip (usually several nanometers to hundreds of nanometers in radius at its apex) mounted at the end of a compliant cantilever beam. In LFM, the tip is held in contact with a sample while it is raster scanned across the sample's surface. The forces that act on the tip are determined by measuring the angular deflections of the cantilever. These forces both map out the sample surface and measure lateral forces. Several techniques can be used to detect the motion of the cantilever; however, the most common method is an optical deflection technique. Here, a laser beam is reflected off the back of the cantilever. The laser beam moves in response to the deflection of the cantilever, and the reflected beam's motion is detected by a four-quadrant position-sensitive photodiode. (See Fig. 2.)

The QCM has been used for decades for microweighing purposes, and was adapted for friction measurements in 1986–1988 by Widom and Krim [21–23]. A QCM consists of a single crystal of quartz that oscillates in transverse shear motion (resonance frequencies are typically at 5–10 MHz) with very little internal dissipation (see Fig. 3 [24]). The oscillations are driven by applying an alternating voltage to thin metal electrodes deposited on the surface of the quartz, the electrodes generally prepared so as to present (111) metal orientations at their surfaces. Atomically thin films later adsorbed onto the QCM electrodes produce shifts in both the frequency and the quality factor Q, which are

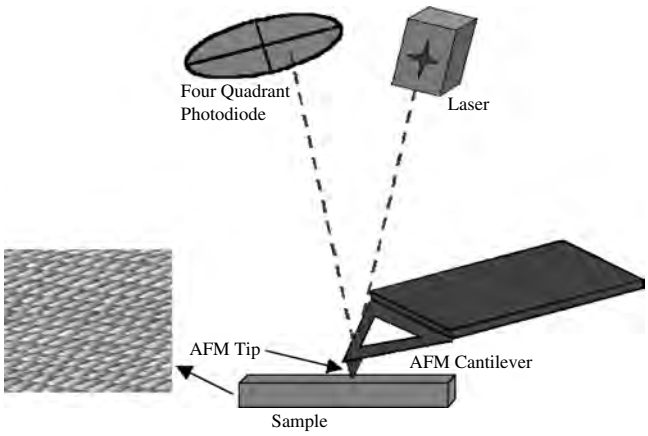


Figure 2. Schematic of the AFM.

indicative of the degree to which the films are able to track the oscillatory motion of the underlying substrate. Friction in this geometry is well characterized by the "viscous friction law" $F/A_c = \eta v$, where v is the average sliding velocity of the adsorbed film constituents. Characteristic slip times τ and friction coefficients (i.e., shear stresses F/A_c per unit velocity) η are determined via the relations [21]

$$\delta(Q^{-1}) = 4\pi\tau(\delta f_o) \quad \eta = \rho_2/\tau \tag{4}$$

where ρ_2 is the mass per unit area of the adsorbate and δf_o is the frequency shift.

Slip times measured by means of QCM are closely linked to the more general topic of energy transfer at an interface, and the vibrational properties of adsorbed molecules [25]. Such vibrations, which underlie heat transfer at an interface, have been studied by means of infrared-reflection-absorption spectroscopy, electron-energy-loss spectroscopy, and inelastic helium scattering. The lifetimes that these techniques determine for the damping of parallel-frustrated translations are of the same order of magnitude, if not identical to, the slip times measured by means of the QCM [26]. The damping rate appears to be little changed in these systems, regardless of the fact that, in the QCM measurements, the vibrational amplitude is larger than several lattice spacings. The fundamental energy dissipative mechanisms involving electronic and phononic effects also appear to be

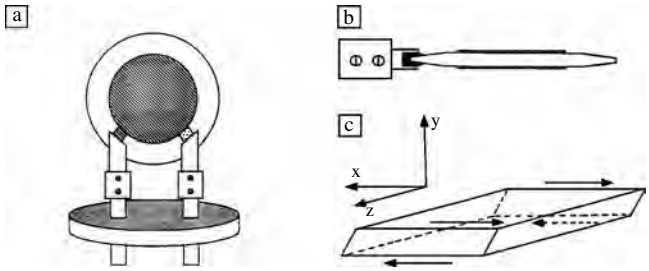


Figure 3. Schematic of the QCM. (a) Front view. The thin quartz disk has metal electrodes deposited onto both faces in a "keyhole" pattern. These electrodes are connected to the leads. (b) Side view. (c) The QCM oscillates in transverse shear mode. Reprinted with permission from [24], M. O. Robbins and J. Krim, *MRS Bull.* 23, 23 (1998). © 1998,

equally viable. Indeed, the QCM slip time is frequently written in terms of separate phonon τ_{ph} and electron–hole τ_{eh} contributions according to $\tau^{-1} = \tau_{ph}^{-1} + \tau_{eh}^{-1}$ [27–32].

The SFA was first developed in the 1950s and 1960s by the Cambridge University group of Bowden and Tabor, and later adapted for friction measurements by Israelachvili in 1973 [18, 19, 33]. The apparatus takes advantage of the fact that the cleavage surface of mica is molecularly smooth, with atomic-step-free areas as great as 1 cm^2. When two mica surfaces are brought into contact, an asperity-free interface is thus formed. The traditional apparatus consists of two cleaved mica surfaces glued to crossed mica cylinders. The contact area and distance between mica surfaces is determined by means of multiple-beam optical interference, with resolution on the order of 0.2 nm or better. The mica surfaces are mounted so that they can be moved vertically and horizontally, and the normal and frictional forces are measured directly from the deflection of a force-mapping spring (see Fig. 4) [34].

4. PHONONIC FRICTION

Tomlinson published an early model, in 1929, of how friction might, in fact, originate at the microscale [35]. In this model, when two surfaces slip across each other, wear-free friction can occur due to the vibration of the atomic lattices. The atoms close to one surface vibrate when the atoms in the opposing surface slip across them. These vibrations are called phonons, or sound waves. The phonons dissipate energy as heat, and this microscopic process is manifested as friction. Motivated by Tomlinson's model, Bowden, Tabor, and colleagues at Cambridge University attempted, in the 1960s, to detect evidence for the phononic contribution to friction, but failed to do so. The idea remained viable, however, as no viable alternate mechanism had presented itself.

Decades later, the concept of phononic friction was independently revisited by Sokoloff et al. in 1978 [36, 37] and McClelland and co-workers in the 1980s, who rederived models similar to the original work of Tomlinson. McClelland, Mate, and colleagues subsequently developed lateral force microscopy (LFM) for studies of atomic-scale friction in an effort to observe phononic mechanisms for friction [20]. To date, the technique has unearthed no direct evidence for phononic mechanisms of friction. It has nonetheless proven to be a revolutionary new probe for studies of atomic-scale friction in particular, and nanoscale science in general.

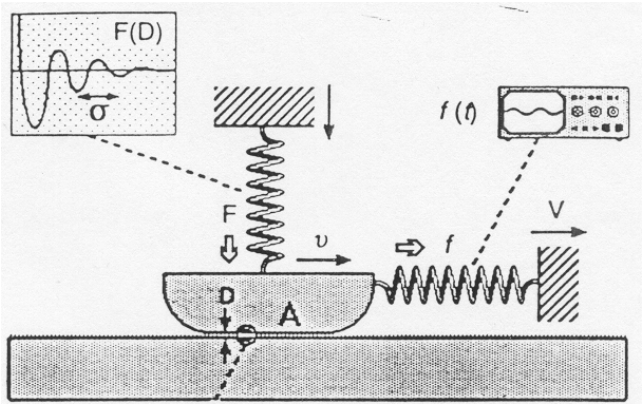


Figure 4. Schematic of the SFA. Reprinted with permission from [34], B. Bhushan et al., *Nature* 374, 607 (1995). © 1995, Macmillan Magazines Ltd.

In 1991, Krim and co-workers reported QCM measurements of the friction of krypton monolayers sliding on Au(111) [38]. They found that solid Kr monolayers sliding on Au(111) had five times longer slip times (lower friction) than liquid monolayers of Kr. In 1994, Cieplak et al. modeled these interesting results [39]. Using standard molecular dynamics algorithms, they reproduced the 1991 experimental results of Krim and colleagues. Their simulations included only one source of friction: phononic friction. Given the close match between their simulations and the QCM experiments, Robbins and colleagues concluded that phononic friction was "the major source of friction" for the krypton on Au(111) system. The combined QCM and numerical results provided the first definitive evidence for the existence of a phononic mechanism for friction.

5. ELECTRONIC FRICTION

In addition to phononic friction, there is another type of dissipation mechanism, electronic, that takes place in electrically conductive materials. In both electronic and phononic friction, kinetic energy associated with the sliding motion at an interface is transformed to heat. The heat is associated with temperature increases in both materials in sliding contact, and is reflected in the phonon populations therein. Phononic dissipative mechanisms involve the direct transfer of energy into the phonon populations. In contrast, electronic mechanisms involve energy transfer into the conduction electrons before transfer to the phonon populations. In one proposed mechanism for electronic friction, the presence of another surface or an adsorbate and the sliding across the metal substrate cause excitations of the conduction electrons near the interface. The excitations generate electron–hole pairs. Electron–hole pairs in metals (in contrast to semiconductors) have very brief lifetimes. When these electron–hole pairs relax, the excess energy is dissipated as heat, and the temperature concomitantly rises. The mechanism is quite distinct from the forces associated with static charge build up on insulating materials.

Persson has suggested that the electronic component of friction μ_{eh} is related to the surface resistivity of an adsorbate–substrate system [26] via

$$\mu_{eh} = n^2 e^2 d \Delta\rho / m n_a \tag{5}$$

Here, n is the number of conduction electrons per unit volume, e is the electron charge, d is the thickness of the metal film, $\Delta\rho$ is the adsorbate-induced increase in film resistivity, m is the adsorbate mass, and n_a is the number of adsorbates per unit area in direct contact with the adsorbate. This is derived for a viscous friction law of the following form:

$$F_{fric} = m \mu_{eh} v \tag{6}$$

where F_{fric} is the force of friction and v is the sliding speed. Definitive proof of electronic contributions to friction is of

great interest since, to date, the vast majority of fundamental theoretical treatments of friction have considered phonon contributions only [40].

Schumacher [41] employed Eq. (5) to infer friction coefficients for a range of adsorbates, and found the slip times to be in the range 10^{-9}–10^{-12} s. The values compared favorably with slip times measured by means of QCM, which probes the total friction $\tau^{-1} = \tau_{ph}^{-1} + \tau_{eh}^{-1}$ [27–32]. In an attempt to quantify the relative contributions of electronic friction to phononic friction, Dayo and Krim compared the electronic friction calculated from electrical resistivity measurements to the total friction taken from QCM measurements for xenon on Ag(111) at 77 K [42]. They observed the resistivity of the Ag(111) film to increase monotonically with the adsorption of the Xe, up to coverages of one monolayer. Above one monolayer, the resistivity increased only very slightly. However, the friction force per unit area required to slide the two-layer thick Xe film was 27% greater than that required to slide the monolayer film [43]. Using Eq. (5), they calculated the electronic contribution to friction, and compared it to the total friction obtained from the QCM, and reported the electronic contribution to be at most 30% of the total for Xe/Ag(111).

There have been several efforts to model the sliding of Xe/Ag(111). In some of these efforts, the models indicate that phononic friction dominates. Tomassone and colleagues [44] reported that the increased number of possible vibrational modes that dissipate energy is greater for a bilayer than a monolayer, and that this is the cause of the increase in friction. In models of the Xe/Ag(111) system by Persson and Nitzan [30], however, the increase in friction from the monolayer to the bilayer was reported to be caused by an increase in the electronic friction. They reported that the second layer pushes the xenon closer to the silver surface, thus increasing the van der Waals contribution to the electronic friction.

Arguably, more direct evidence supporting the importance of the electronic contributions comes from the work published by Dayo et al. [45] in 1998. The group studied the sliding of thin nitrogen films adsorbed on lead surfaces by means of QCM, and observed the friction coefficient to drop by almost half when the substrate became superconducting. The result drew much attention, and sparked criticism as well [46–61]. A somewhat similar experiment found no slippage whatsoever of a nitrogen film at low temperature, and thus no changes in the friction across the superconducting transition [58]. It subsequently became clear, however, that the effect is observable only on substrates of extremely high purity: pinning of the film at low temperature by contamination or defects prevails for all but the cleanest substrates [62, 63]. The experiment remains theoretically puzzling. Indeed, various arguments have been presented that justify, contradict, or partially explain the observation [51–61].

Tribologists are not the only research community to have pondered energy dissipation mechanisms at moving interfaces. An entirely distinct surface science community emerged in the late 1980s, members of which studied both phononic and electronic contributions to energy dissipation within the context of small vibrational motions of atoms on surfaces. Whenever atoms or molecules adsorb on surfaces, new vibrational modes will emerge which are not present in either an isolated surface or the adsorbate alone. The modes which appear include both "internal" (stretching or torsional vibrations within a molecule) and "external" modes, whereby the entire molecule or atom moves as a whole with respect to the surface. These phonon modes for adsorbate–substrate systems can be measured by helium atom scattering (HAS) and related spectroscopies [64, 65]. They can also measure the line width (the lifetime or damping) of adsorbate vibrations, which is indicative of the rate at which the energy of the phonon mode is dissipated into heat. The line width of the adsorbate vibration has been shown to be linearly proportional to the friction coefficient η [66]. Frustrated translational modes, both parallel and perpendicular to the substrate, have been measured. Perpendicular modes (S modes or FT_z modes) mainly probe the energy and damping of the interaction between the adsorbate and substrate. Parallel modes (T modes) probe the curvature of the adsorbate–substrate interaction potential [67–69] and the commensurability between the adsorbate and substrate [69].

From the information given by the phonon dispersion curves and the line width of the adsorbate vibrations gleaned from HAS and related spectroscopies, it should be possible to predict and explain the friction for a given adsorbate–substrate system. For example, since the line width of the adsorbate vibrations is linearly proportional to the friction coefficient, a broader line width should imply a higher friction system. Also, the existence of a Brillouin-zone-center gap for parallel or perpendicular phonon modes indicates commensurability between the adsorbate and substrate [69]. The existence and size of the zone-center gap may therefore predict the degree of lattice match up that causes friction. This understanding, however, is still in its infancy. Although the comparative magnitude of friction between adsorbates and substrates may be predictable, it is still difficult to determine whether the origins of friction for a given system are phononic or electronic. Further well-planned and controlled experiments are necessary to differentiate between phononic and electronic processes. Of particular interest may be the QCM experiments by Mistura and colleagues [62, 70], whereby QCM vibrational amplitudes are sufficiently small so as to be directly comparable to the vibrational sliding distances encountered in the "vibrations at surfaces" community.

6. STATIC FRICTION

One of the most common everyday experiences with friction at the macroscopic scale is the occurrence of static friction: the force to initiate motion (which itself is quite variable, depending, for example, on how long the two surfaces have been in contact) is virtually always larger than that required to keep an object in motion. A closely associated phenomenon is that of stick-slip friction, whereby, for certain sliding speeds, the velocity-weakening dependence of the transition from static to sliding friction leads to repetitive sticking and slipping at the interface, producing the all-too-familiar screeching noises associated with brakes. Although static friction is ubiquitous, it is notoriously difficult to explain at the macroscopic and microscopic length scales.

Static friction and stick-slip motion are also familiar occurrences in nanotribology. Static friction has frequently been measured in the SFA [71, 72]. Atomic-scale stick-slip friction was first measured in the AFM by Mate and colleagues in 1987 [20]. Using an AFM with a tungsten probe on a graphite surface, they found that the frictional force of the probe on the graphite had a periodicity of 2.5 Å, the same periodicity as the graphite surface. Since 1987, atomic-scale stick-slip friction has been observed on many different materials [73–75]. Atomic-scale stick-slip friction can be interpreted as the tip and sample remaining in the minimum potential energy position until a sufficiently strong shear stress is applied to force the tip to move [76]. Interestingly, such stick-slip behavior has always been observed to occur with the periodicity of the substrate lattice—even when the lattice points have multiple atoms (Fig. 5). For example, it is not possible to differentiate between the molybdenum and the sulfur atoms in an MoS_2 surface. The stick-slip periodicity measured by the LFM is 3.16 Å, which corresponds to the distance between MoS_2 molecules in the lattice [77, 78].

Static friction has never been evident in QCM measurements, both solid–solid and liquid–solid interfaces being well described by the viscous friction law in Eq. (1). Given the differences in sliding speeds and geometries among QCM, SFA, and AFM, there was some debate about the underlying reason for the lack of static friction in the QCM. A recent experiment by Mate and Marchon explored this question [79]. In this experiment, they focused on the fact that, while both QCM and SFA measurements of the shearing of liquid films reveal viscous friction, static friction is present only in the SFA geometry. Mate and Marchon, employing a "blow-off experiment," explored whether the open geometry or much greater shear rates of the QCM could account for the difference in the observed behaviors. They blew nitrogen across a liquid film confined in a narrow channel. The laminar flow conditions generated a shear stress on the liquid film at shear rates comparable to those in the SFA. Their results yielded viscous friction, so it was concluded that confined geometries, and not the shear rates, of the SFA and AFM resulted in static friction, while unconfined geometries like the QCM and the "blow-off" experiment are characterized by the viscous friction law.

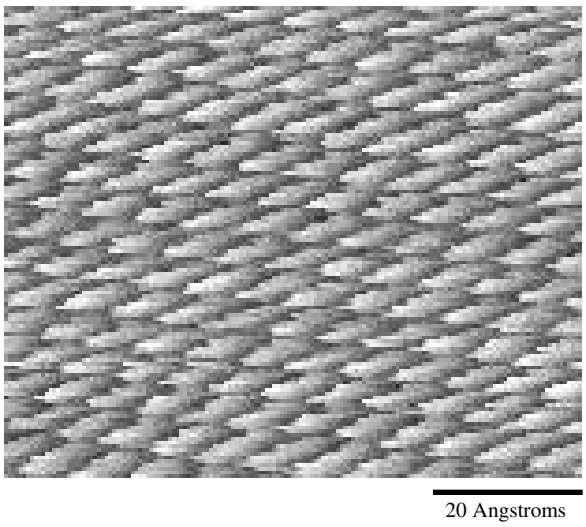


Figure 5. AFM friction force image of atomic stick-slip friction. Acquired in Krim lab with Si_3N_4 tip on muscovite mica. Each sticking point corresponds to one lattice point on the mica surface.

The results are consistent with recent computer simulation experiments [80, 81]. Robbins and colleagues studied the effects that mobile atoms between two contacting crystalline surfaces have on friction. As discussed previously, commensurate crystalline surfaces are expected to have high friction, and they exhibit static friction with or without contaminant atoms between the surfaces. Incommensurate crystalline surfaces, however, have very low friction. In the real world, however, almost all contacting surfaces are expected to be incommensurate, so the existence of static friction is difficult to explain. In their simulations, Robbins and colleagues modeled submonolayer, mobile contaminant films between two incommensurate surfaces. They observed that the mobile molecules found small gaps between the surfaces. Any sliding of the two surfaces will constrict or reduce the gaps in which the mobile molecules settle. The molecules therefore resist sliding motion until the shear stress exceeds some threshold value and they are displaced, resulting in the presence of static friction.

7. VELOCITY DEPENDENCE OF FRICTION

Recent LFM experiments have shown that the velocity independence of friction for multiasperity contacts and macroscale objects does not hold for single-asperity contacts and nanoscale objects [82]. Gnecco and colleagues [83] studied the velocity dependence of friction by means of LFM for a silicon tip sliding on an NaCl(100) surface in ultrahigh vacuum (UHV) for speeds ranging from 5 nm/s to 1 μm/s. They observed the atomic stick-slip friction common to LFM, and also observed that the friction increased logarithmically with the sliding velocity. Employing a modified Tomlinson model, they found that the velocity dependence was attributable to thermal energy that helped to push the AFM tip from one lattice "sticking point" to the next.

In contrast to the logarithmic velocity dependence associated with LFM (a confined geometry), friction data obtained using the open geometry of the QCM are routinely described by the viscous friction law $F = \eta v A_c$. Mak and Krim confirmed the applicability of this law for monolayer coverages of krypton adsorbed on Au(111) at 77 K, for sliding speeds in the range 3–8 cm/s [84]. In contrast to full monolayer coverage, the frictional law appeared to be nonlinear for submonolayer coverages, described by the more general law v^n, with $n < 1$. Numerical simulations have also been performed on the Kr/Au(111) system, indicating a linear friction law for sliding speeds on the order of 10 m/s [39]. The open geometry of Mate and Marchon's "blow-off" experiment, mentioned above [79], also obeyed the viscous friction law, but for shear rates nine orders of magnitude slower.

8. FRICTIONAL ANISOTROPY

There has been much work done on commensurate and incommensurate frictional systems and frictional anisotropy. When two surfaces are commensurate, their lattices match in perfect registry. Theoretically, it has been shown that commensurate systems can have as much as 14 orders of magnitude higher friction than corresponding incommensurate systems. The lower friction levels for incommensurate systems are due to the random positions of atoms with respect to each other on the mismatched lattices. Interfacial forces tend to cancel in this case, resulting in very low friction [85, 86].

There are several experiments that reveal the effect of incommensurability on friction, although none has reported the phenomenal change in friction that is theoretically possible. Most of these studies study the adhesion and friction between like materials or materials with similar lattice spacing, and go from commensurate to incommensurate by rotating the lattices with respect to one another (see Fig. 6). McGuiggan and Israelachvili, for example, employed an SFA to study the adhesion of mica surfaces in distilled water and aqueous KCl [87]. Their results related the adhesion of the mica surfaces to the alignment of the mica surfaces with respect to one another. The adhesion peaked at $\theta =$ 0, 60, 120, and 180° ($\theta = 0$ corresponds to matching lattices), following the sixfold symmetry of the mica lattice. A related study of mica surfaces by Hirano et al. linked the friction of the mica surfaces to their relative orientation [88]. They used a friction tester similar to an SFA, and studied the rubbing of two mica surfaces in both an argon-purged dry atmosphere with the surface temperature of the mica heated above 100 °C to prevent water adsorption. These measurements were then compared to ambient-condition measurements with the mica surface temperatures held at 20 °C. In the argon-purged atmosphere, the friction reached maximum levels when the surfaces were commensurate ($\theta = 0$, 60, 120, and 180°, $\theta = 0$ corresponds to matching lattices) and minimum levels, dropping by a factor of 4 when the surfaces were incommensurate ($\theta = 30$, 90, etc.). In ambient conditions, the frictional anisotropy disappeared, however, owing to the adsorption of water and other contaminants onto the mica surfaces. The result lent support to the idea that frictional anisotropy is purely a surface effect. Hirano and colleagues [89] later employed scanning tunneling microscopy (STM) in UHV to study the sliding of a tungsten wire tip surface on a planar silicon surface. Much like an AFM, they rastered the tungsten tip across the silicon surface, and used a displacement sensor to measure the bending of the tungsten wire caused by the friction between the two surfaces. Employing Hooke's law to extract the frictional force from the displacement of the tungsten tip, they found that the frictional force of the two surfaces in commensurate contact was 8×10^{-8} N. The friction vanished when the surfaces were in incommensurate contact, within their resolution. (The resolution of their instrument was 3×10^{-9} N.)

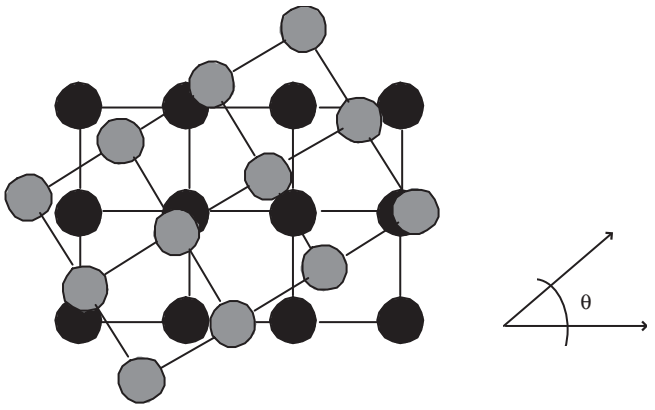


Figure 6. Two identical lattices can be made incommensurate by rotating one lattice by θ and sliding along the new direction.

Not all frictional anisotropy can be related to surface commensurability. In some experiments, frictional anisotropy disappears at low loads when the surfaces are in elastic contact, and is only present at high loads when the surfaces are in plastic contact [90]. Take, for example, an experiment by Ko and Gellman [91] that studied the friction between nickel(100) surfaces in UHV with and without adsorbed surface coatings of sulfur and ethanol. The adsorbed coatings modified the surface lattice from the bare nickel: the sulfur surface is well ordered, with lattice vectors rotated 45° with respect to the bulk nickel, and the adsorbed ethanol is not ordered. The friction in this study was anisotropic, and related to the bulk lattice structure of the nickel. The static friction coefficient was found to be minimized (lower by a factor of 3–4) at a bulk lattice mismatch of $\theta = 45°$, regardless of the modifications to the surface by monolayers of atomic sulfur and up to four monolayers of adsorbed ethanol. Their results suggest that, in some systems in plastic contact, the frictional anisotropy is related to the deformation of the bulk material.

9. CONCLUSIONS

Since the development in the 1970s and 1980s of important new technology that enabled the study of nanotribology, much has been learned. However, nanotribology is a new field with much left to be discovered. There is still much to be learned about the chemical reactions which occur in a sliding contact, the characterization of the microstructural and mechanical properties of the contact regions between sliding materials, bridging the gap between nanotribology and macrotribology, developing realistic computer simulations, explaining the origins of friction, and establishing realistic laboratory test setups which are both well controlled and relevant to operating machinery.

GLOSSARY

Amontons' law At macroscopic length scales, the friction is proportional to the load.

Commensurate When two surface lattices are in perfect registry.

Coulomb's friction law At macroscopic length scales, friction force is independent of sliding velocity.

Electronic friction An energy dissipation mechanism that occurs in electrically conductive materials. During sliding,

energy is transferred to the conduction electrons prior to being transferred to phonons, as in phononic friction.

Frictional anisotropy The fact that frictional forces are dependent on alignment of the surface lattices.

Incommensurate When two surface lattices are not in perfect registry.

Nanotribology The study of friction, lubrication, and wear at atomistic length and time scales.

Phonon The mechanical vibration of an atomic lattice, or sound wave.

Phononic friction The process occurring when the atoms close to one surface vibrate, generating phonons, when the atoms in the opposing surface slip across them. The phonons then dissipate energy as heat, causing friction.

Static friction The force required to initiate motion.

Stick-slip friction The velocity-weakening dependence of the transition from static to sliding friction that leads to repetitive sticking and slipping at the interface.

Tribology The study of friction, lubrication, and wear.

REFERENCES

1. H. P. Jost, *Wear* 136, 1 (1990).
2. K. C. Ludmema, "Friction, Wear, Lubrication: A Textbook in Tribology." CRC Press, Boca Raton, FL, 1996.
3. M. Abdelmaksoud, J. Bender, and J. Krim, *Phys. Rev. Lett.* (submitted).
4. B. Borovsky, M. Abdelmaksoud, and J. Krim, in "Nanotribology: Critical Assessment and Research Needs" (S. M. Hsu and Z. C. Ying, Eds.), pp. 361–375. Kluwer, Boston, 2002.
5. M. T. Dugger, in "Nanotribology: Critical Assessment and Research Needs" (S. M. Hsu and Z. C. Ying, Eds.), pp. 123–137. Kluwer, Boston, 2002.
6. M. Porto, M. Urbach, and J. Klafter, *Phys. Rev. Lett.* 84, 1608 (2000).
7. G. Amontons, *Mem. Acad. Roy. Sci.* 206 (1699).
8. F. Palmer, *Sci. Amer.* 184 (1951).
9. R. W. Carpick, N. Agrait, D. F. Ogletree, and M. Salmeron, *J. Vac. Sci. Technol. B* 14, 1289 (1996).
10. M. A. Lantz, S. J. O'Shea, M. E. Welland, and K. L. Johnson, *Phys. Rev. B* 55, 10776 (1997).
11. M. A. Lantz, S. J. O'Shea, and M. E. Welland, *Phys. Rev. B* 56, 15345 (1997).
12. M. Enachescu, R. J. A. van den Oetelaar, R. W. Carpick, D. F. Ogletree, C. F. J. Flipse, and M. Salmeron, *Phys. Rev. Lett.* 81, 1877 (1998).
13. M. Enachescu, R. J. A. van den Oetelaar, R. W. Carpick, D. F. Ogletree, C. F. J. Flipse, and M. Salmeron, *Tribology Lett.* 7, 73 (1999).
14. K. L. Johnson, K. Kendall, and A. D. Roberts, *Proc. R. Soc. London A* 324, 301 (1971).
15. B. V. Derjaguin, V. M. Muller, and Y. P. Toporov, *J. Colloid Interface Sci.* 53, 314 (1975).
16. H. Hertz, in "Gessamelte Werke" (P. Lenard and J. A. Barth, Eds.). Leipzig, 1895.
17. F. P. D. Bowden and D. Tabor, "Friction and Lubrication of Solids, Part I." Oxford University Press, Oxford, 1954.
18. J. N. Israelachvili and D. Tabor, *Proc. R. Soc. London A* 331, 19 (1972).
19. J. N. Israelachvili and D. Tabor, *Prog. Surf. Membr. Sci.* 7, 1 (1973).
20. C. M. Mate, G. M. McClelland, R. Erlandsson, and S. Chang, *Phys. Rev. Lett.* 59, 1942 (1987).
21. J. Krim and A. Widom, *Phys. Rev. B* 38, 12184 (1986).
22. A. Widom and J. Krim, *Phys. Rev. B* 34, R3 (1986).
23. E. T. Watts, J. Krim, and A. Widom, *Phys. Rev. B* 41, 3466 (1990).
24. M. O. Robbins and J. Krim, *MRS Bull.* 23, 23 (1998).
25. A. Zangwill, "Physics at Surfaces," p. 328. Cambridge University Press, Cambridge, 1988.
26. B. N. J. Persson, *Phys. Rev. B* 44, 3277 (1991).
27. B. N. J. Persson and E. Tosatti, Eds., "Physics of Sliding Friction." Kluwer, Dordrecht, 1996.
28. B. N. J. Persson, "Sliding Friction." Springer-Verlag, New York, 1997.
29. B. N. J. Persson, *Surf. Sci. Rep.* 33, 85 (1999).
30. B. N. J. Persson and A. Nitzan, *Surf. Sci.* 367, 261 (1996).
31. A. Liebsch, S. Goncalves, and M. Kiwi, *Phys. Rev. B* 60, 5034 (1999).
32. E. D. Smith, M. O. Robbins, and M. Cieplak, *Phys. Rev. B* 54, 8252 (1996).
33. D. Tabor and R. H. S. Winterton, *Proc. R. Soc. London A* 312, 435 (1969).
34. B. Bhushan, J. N. Israelachvili, and U. Landman, *Nature* 374, 607 (1995).
35. G. A. Tomlinson, *Phil. Mag.* 7, 905 (1929).
36. J. E. Sacco and J. B. Sokoloff, *Phys. Rev. B* 18, 6549 (1978).
37. J. B. Sokoloff, J. E. Sacco, and J. F. Weisz, *Phys. Rev. Lett.* 41, 1561 (1978).
38. J. Krim, D. H. Solina, and R. Chiarello, *Phys. Rev. Lett.* 66, 181 (1991).
39. M. Cieplak, E. D. Smith, and M. O. Robbins, *Science* 265, 1209 (1994).
40. I. L. Singer and H. M. Pollock, "Fundamentals of Friction: Macroscopic and Microscopic Processes." Kluwer, Dordrecht, 1992.
41. D. Schumacher, "Surface Scattering Experiments with Conduction Electrons," Sect. 4.2–4.3. Springer, Berlin, 1993.
42. A. Dayo and J. Krim, *Int. J. Thermophys.* 19, 827 (1998).
43. C. Daly and J. Krim, *Phys. Rev. Lett.* 76, 803 (1996).
44. M. S. Tomassone, J. B. Sokoloff, and J. Krim, *Phys. Rev. Lett.* 79, 4798 (1997).
45. A. Dayo, W. Alnasrallah, and J. Krim, *Phys. Rev. Lett.* 80, 1690 (1998).
46. Superconductivity-dependent sliding friction, *Phys. News* (1998).
47. Superconductivity and friction, *The Economist* (*Sci. Technol. Sect.*) (Feb. 28, 1998).
48. Superconductivity-dependent friction, *Phys. Today* (*Sci. Technol. Sect.*) (Apr. 1998).
49. E. Stokstad, *Science NOW*, Website of *Science Mag.* (Feb. 23, 1998).
50. I. Peterson, *Sci. News* (*Phys. Res. Notes*) (Mar. 14, 1998).
51. B. N. J. Persson and E. Tosatti, *Surf. Sci.* 411, L855 (1998).
52. T. Novotny and B. Velicky, *Phys. Rev. Lett.* 83, 4112 (1999).
53. J. B. Sokoloff, M. S. Tomassone, and A. Widom, *Phys. Rev. Lett.* 84, 515 (2000).
54. L. W. Bruch, *Phys. Rev. B* 61, 16201 (2000).
55. B. N. J. Persson and Z.Y. Zhang, *Phys. Rev. B* 57, 7327 (1998).
56. J. B. Sokoloff, *Phys. Rev. B* 52, 5318 (1995).
57. V. L. Popov, *Phys. Rev. Lett.* 83, 1632 (1999).
58. R. L. Renner, J. E. Rutledge, and P. Taborek, *Phys. Rev. Lett.* 83, 1261 (1999).
59. V. L. Popov, *JETP Lett.* 69, 558 (1999).
60. A. I. Volokitin and B. N. J. Persson, *Phys. Low-Dimens. Str.* 7–8, 17 (1998).
61. B. N. J. Persson, "Sliding Friction," Chap. 8. Springer-Verlag, New York, 1997.
62. L. Bruschi, A. Carlin, and G. Mistura, *Phys. Rev. Lett.* 88, 046105 (2002).
63. R. L. Renner, J. E. Rutledge, and P. Taborek, *Phys. Rev. B* 63, 233405 (2001).
64. J. P. Toennies, *J. Phys. Cond. Matt.* 5, A25 (1993).

65. Ch. Woll, in "Physics of Sliding Friction" (B. N. J. Persson and E. Tosatti, Eds.), p. 231. Kluwer, Dordrecht, 1996.
66. B. N. J. Persson, "Sliding Friction, Physical Principles and Applications." Springer, Berlin, 2000.
67. G. Witte, *Surf. Sci.* 502–503, 405 (2002).
68. J. Braun, D. Fuhrmann, A. Siber, B. Gumhalter, and Ch. Woll, *Phys. Rev. Lett.* 80, 125 (1998).
69. L. W. Bruch and F. Y. Hansen, *Phys. Rev. B* 55, 1782 (1997).
70. L. Bruschi and G. Mistura, *Phys. Rev. B* 63, 235411 (2001).
71. A. L. Demirel and S. Granick, *J. Chem. Phys.* 109, 6889 (1998).
72. J. N. Israelachvili, Y. L. Chen, and H. Yoshizawa, *J. Adhes. Sci. Technol.* 8, 1231 (1994).
73. H. Takano and M. Fujihira, *J. Vac. Sci. Technol. B* 14, 1272 (1996).
74. G. J. Germann, S. R. Cohen, G. Neubauer, G. M. McClelland, H. Seki, and D. Coulman, *J. Appl. Phys.* 73, 163 (1993).
75. J. Ruan and B. Bhushan, *J. Appl. Phys.* 76, 5022 (1994).
76. R. W. Carpick and M. Salmeron, *Chem. Rev.* 97, 1163 (1997).
77. S. Fujisawa, E. Kishi, Y. Sugawara, and S. Morita, *Jpn. J. Appl. Phys.* 33, 3752 (1994).
78. S. Morita, S. Fujisawa, and Y. Sugawara, *Surf. Sci. Rep.* 23, 1 (1996).
79. C. M. Mate and B. Marchon, *Phys. Rev. Lett.* 85, 3902 (2000).
80. G. He, M. H. Muser, and M. O. Robbins, *Science* 284, 1650 (1999).
81. M. H. Muser and M. O. Robbins, *Phys. Rev. B* 61, 2335 (2000).
82. T. Bouhacina, J. P. Aime, S. Gauthier, D. Michel, and V. Heroguez, *Phys. Rev. B* 56, 7694 (1997).
83. E. Gnecco, R. Bennewitz, T. Gyalog, Ch. Loppacher, M. Bammerlin, E. Meyer, and H. J. Guntherodt, *Phys. Rev. Lett.* 84, 1172 (2000).
84. C. Mak and J. Krim, *Phys. Rev. B* 58, 5157 (1998).
85. M. Hirano and K. Shinjo, *Phys. Rev. B* 41, 11837 (1990).
86. K. Shinjo and M. Hirano, *Surf. Sci.* 283, 473 (1993).
87. P. M. McGuiggan and J. N. Israelachvili, *J. Mater. Res.* 5, 2332 (1990).
88. M. Hirano, K. Shinjo, R. Kaneko, and Y. Murata, *Phys. Rev. Lett.* 67, 2642 (1991).
89. M. Hirano, K. Shinjo, R. Kaneko, and Y. Murata, *Phys. Rev. Lett.* 78, 1448 (1997).
90. Y. Enonmoto and D. Tabor, *Proc. R. Soc. London A* 373, 405 (1981).
91. J. S. Ko and A. J. Gellman, *Langmuir* 16, 8343 (2000).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanotribology: Friction Force Microscopy

Enrico Gnecco
University of Basel, Basel, Switzerland

CONTENTS

1. INTRODUCTION

The term *tribology* comes from the ancient Greek word *tribos*, which means *to rub*. It was used for the first time in 1966 by Peter Jost, and, in current use, it refers to the study of friction, wear, lubrication, and contact mechanics. *Nanotribology* encompasses the same topics, but it focuses on processes occurring on the nanometer scale. Nanotribology aims to realize two important tasks: (i) to give insight into the microscopic origin of friction and wear and (ii) to exploit the knowledge acquired for the optimization of microdevices. The economic impact of nanotribology is easily imaginable. Significant parts of gross domestic product of industrialized countries are wasted every year due to friction and wear problems which, even if they affect macroscopic machinery, have their origin in complex processes of various natures occurring on the nanometer scale. These processes are even more relevant in data storage and microelectronics industry. The need for high recording densities in magnetic storage devices requires a reduction in the height of the read/write magnetic head with respect to the magnetic medium, which can lead to contact with serious consequences. In microelectromechanical systems, objects a few micrometers in size are repeatedly brought into contact at high speed. In both cases the resistance to rupture of different coatings and components needs to be tested and controlled on the small scale.

The macroscopic results of tribology cannot be directly extrapolated to the micro- and nanoworld. If the linear size of a component is reduced by a factor of 10, the area of its surface is reduced 10 times less than its volume, which implies that the resistive forces, proportional to the area of sliding, are reduced 10 times less than the inertial forces, proportional to the volume of the component. The observation of tribological processes down to the nanometer scale was made possible by the *atomic force microscope* (AFM), introduced by Binnig et al. in 1986 [1]. The AFM reveals the forces acting on a sharp tip sliding on a surface from the deformation of a microcantilever supporting the tip (Fig. 1). The normal force between tip and surface causes the vertical deflection of the cantilever, whereas the friction force causes its torsion. If the vertical deflection is kept constant by a feedback loop, friction can be monitored at a given load while scanning. In such context, the AFM is also referred as the *friction force microscope* (FFM). How to accurately detect friction with AFM is the main subject of Section 2. Other instruments of interest in nanotribology—the surface force apparatus and the quartz crystal microbalance—will be also briefly introduced.

The contact between two macroscopic surfaces involves thousands of microasperities. The apex of the AFM tip sliding on the underlying sample represents a single asperity a few nanometers in size. The mechanics of a single contact differs from that of a multiple contact. If the asperity has a spherical shape the friction or lateral force, F_L, is related to the applied normal force, F_N, by the Hertzian relation $F_L \propto F_N^{2/3}$ [3], which seems to be in contrast with the well-known *Amonton's law*, $F_L \propto F_N$ [4]. However, Amonton's law is recovered when statistical distributions of single asperities are considered [5, 6]. Adhesion plays an important role in the contact region. This effect is considered in extended Hertzian models, which well fit the experiments conducted with the FFM. These models, all based on continuum mechanics, will be discussed in Section 3.

In a well-defined environment, the friction force acting on the probing tip depends on the surface under investigation. Due to the small size of the tip (typical radii of curvature are 10–50 nm) AFMs can map forces on a local scale with extraordinary resolution. An example is given in Figure 2,

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (879–898)

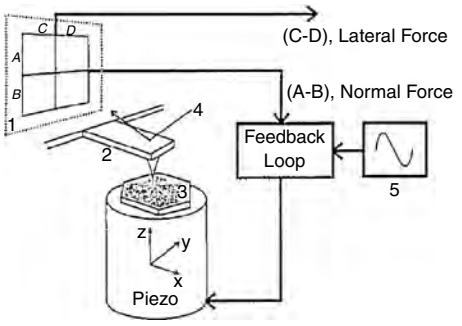


Figure 1. Schematic diagram of a beam-deflection friction force microscope. (1) Four-quadrant photodiode; (2) cantilever with integrated probing tip; (3) segmented piezoelectric tube with sample; (4) ray trace of the laser beam. Reprinted with permission from [2], R. Lüthi et al., *Surf. Sci.* 338, 247 (1995). © 1995, Elsevier Science.

which shows a friction map acquired on some hydrocarbon islands surrounded by fluorocarbons. Using a high load it was possible to produce patterns on the hydrocarbon islands, whereas fluorocarbons could not be affected by this treatment. Further examples of friction maps on the nanometer scale are given in Section 4.

Friction maps with atomic features were first detected by Mate et al. on graphite in 1987 [8]. Figure 3a shows a friction map of sodium chloride acquired in ultra-high-vacuum (UHV) conditions, which reproduces the structure of the atomic lattice. A single scan line has a sawtooth shape with the periodicity of the lattice (Fig. 3b). When the scan direction is inverted the lateral force is also inverted. The area of the *friction loop* obtained by scanning backward and forward gives the energy dissipated per scan cycle. The sawtooth shape is due to the *atomic stick-slip* movement of the probing tip, which, driven by the cantilever, jumps periodically from a stable equilibrium position on the lattice to the next one. A simple explanation of the atomic stick-slip is given by the *Tomlinson model* [10], discussed in Section 5. An extension of the Tomlinson model, which includes thermal effects, predicts a logarithmic dependence of friction on the sliding speed, in contrast with the classical *Coulomb's law of friction*,

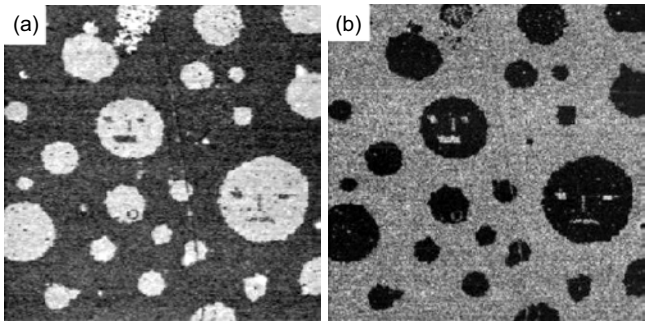


Figure 2. (a) Topography of mixed Langmuir–Blodgett film. Areas of hydrocarbons are surrounded by a "sea" of fluorocarbons. (b) Friction map. Small areas could be removed by the probing tip within the hydrocarbon islands, and not within fluorocarbons. Reprinted with permission from [7], E. Meyer et al., *Thin Solid Films* 220, 132 (1992). © 1992, Elsevier Science.

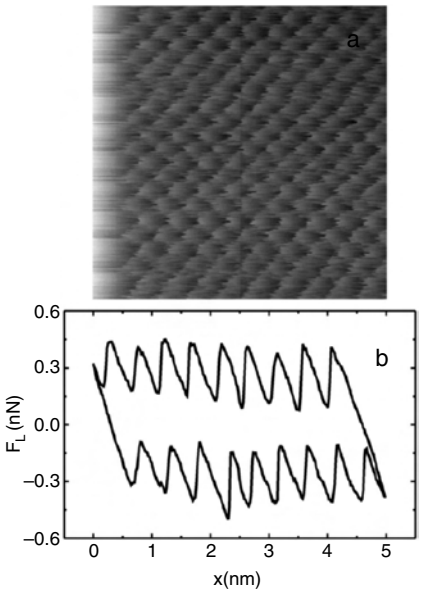


Figure 3. (a) Friction map of sodium chloride. (b) Friction loop acquired in a forward and backward scan. Reprinted with permission from [9], E. Gnecco et al., *Phys. Rev. Lett.* 84, 1172 (2000). © 2000, American Physical Society.

which states that friction is speed-independent [4]. Section 6 deals with friction measurements on the atomic scale and related phenomena, as friction anisotropy, dissipative effects "out of contact," and electronic friction.

If the normal load applied on the probing tip exceeds a critical value, the morphology of the surface is permanently modified. *Nanoindentation* can be used to quantify hardness on the nanoscale. In the *nanolithography* technique the AFM is used to produce nanopatterns in a controlled way. In exceptional cases wear processes can be studied down to the atomic scale. For example, Figure 4 shows that the debris extracted from an ionic crystal recrystallized in mounds with the same structure of the undamaged surface. Tip-induced modifications of various surfaces will be presented in Section 7.

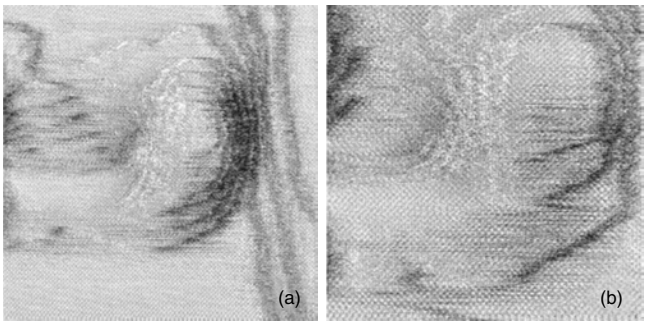


Figure 4. Lateral force images acquired at the end of a groove scratched on potassium bromide. The debris recrystallizes in a mound with ordered structure. Frame sizes: (a) 39 nm, (b) 25 nm. Adapted with permission from [11], E. Gnecco et al., *Phys. Rev. Lett.* 88, 215501 (2002). © 2002, American Physical Society.

2. INSTRUMENTS

2.1. Friction Force Microscopy

The atomic force microscope exploits the interaction of a sharp tip sliding on a surface to quantify dissipative processes down to the atomic scale (Fig. 1). The relative motion of tip and surface is realized by a scanner formed by piezoelectric elements, which moves the surface perpendicularly to the tip with a certain periodicity. The scanner can be also extended or retracted in order to vary the normal force, F_N, applied to the surface. This force causes the vertical deflection of the cantilever supporting the tip. If the normal force increases while scanning, due to the local slope of the surface, the scanner is retracted by a feedback loop to keep the force F_N constant. On the other hand, if F_N decreases, the surface is brought closer to the tip by extending the scanner. In such a way, the topography of the surface can be determined line by line. An accurate control of the piezo movement is made possible by a laser beam, which is reflected from the rear of the cantilever into a photodetector. When the bending of the cantilever changes, the laser spot on the detector moves up or down. The difference signal between upper and lower segments $(A - B)$ is proportional to the normal force.

The relative sliding of tip and surface is also accompanied by *friction*. A lateral force, F_L, with the opposite direction of the scan velocity, hinders the motion of the tip. The friction force provokes the torsion of the cantilever and it can be observed with the topography if the photodetector reveals not only the vertical deflection but also the lateral movement of the cantilever while scanning. In practice, this is realized by four-quadrant photodetectors, where the difference signal between left and right segments $(C - D)$ is proportional to the friction force. Incidentally, friction causes also the lateral bending of the cantilever, but this effect is negligible if the thickness of the cantilever is much smaller than its width.

The AFM was first used by Mate et al. to reveal friction with atomic features [8]. In their experiment, Mate et al. used a tungsten wire and a slightly different technique to reveal the forces acting in the contact area. The optical beam deflection was introduced later by Marti et al. [12] and Meyer and Amer [13]. Other methods to measure the forces between tip and surface are based on capacitance detection [14], dual fiber interferometry [15], and piezolevers [16]. In the first method, two plates close to the cantilever reveal capacitance while scanning. The second technique uses two optical fibers to detect the cantilever deflection along two orthogonal directions angled 45° with the respect to the surface normal. In the third method, cantilevers with two Wheatstone bridges at their base are used. Normal and lateral forces are respectively proportional to the sum and the difference of both bridge signals.

2.2. Force Calibration in Friction Force Microscopy

Force calibration is relatively simple for rectangular cantilevers [2]. Due to possible discrepancies with the geometric values provided by manufacturers, it is recommended to use optical and electron microscopes to determine the width w, the thickness t, and the length l of the lever, as well as the tip height, h, and the position of the tip with respect to the cantilever. The cantilever thickness can be also deduced accurately from the resonance frequency, f_0, using the relation [17]

$$t = \frac{2\sqrt{12}\pi}{1.875^2}\sqrt{\frac{\rho}{E}}f_0 l^2 \tag{1}$$

In (1) ρ is the cantilever density, and E is its Young's modulus. The normal spring constant, c_N, and the lateral spring constant, c_L, are given by

$$c_N = \frac{Ewt^3}{4l^3} \qquad c_L = \frac{Gwt^3}{3h^2 l} \tag{2}$$

where G is the shear modulus of the cantilever. The further step consists in measuring the sensitivity of the photodetector, S_z (nm/V). The sensitivity of a beam-deflection AFM is determined by force vs distance curves measured on hard surfaces, where elastic deformations are negligible and the vertical movement of the scanner equals the cantilever deflection. A typical relation between the difference of the vertical signals on the four-quadrant detector, V_N, and the distance from the surface, z, is sketched in Figure 5. When the tip is approached no signal is revealed until the tip jumps into contact at $z = z_1$. Further extension or retraction of the scanner result in an elastic response until the tip jumps again out of contact at a distance $z_2 > z_1$, due to adhesive forces. The slope of the elastic part of the curve gives the required sensitivity, S_z. Assuming that the laser beam is well positioned above the probing tip, the normal and lateral forces are related to the voltage V_N, and the difference between the horizontal signals, V_L, on the photodetector by the following relations:

$$F_N = c_N S_z V_N \qquad F_L = \frac{3}{2}c_L \frac{h}{l} S_z V_L \tag{3}$$

The normal spring constant c_N can be evaluated with alternative methods. Cleveland et al. attached tungsten spheres to the tip, which changes the resonance frequency f_0 according to the formula [18]

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{c_N}{M + m^*}} \tag{4}$$

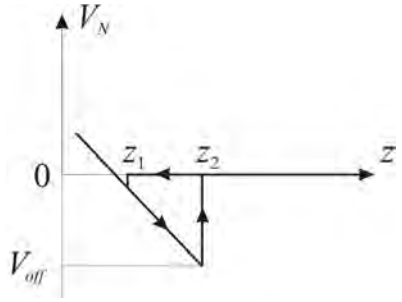


Figure 5. Sketch of a typical force vs distance curve. The tip approaches the sample from right to left. At $z = z_1$ the tip jumps into contact. Further moving results in an elastic bending of the cantilever. Due to adhesive forces, the tip jumps out of contact at $z_2 > z_1$.

In (4) M is the mass of the added object and m^* is a certain effective mass of the cantilever, which depends on its geometry. The spring constant can be deduced from the frequency shifts corresponding to the different masses attached. Hutter and Bechhoefer [19] observed that the constant c_N is related to the area of the power spectrum of the thermal fluctuations of the cantilever, P. The correct relation is $c_N = 4k_BT/3P$, where $k_B = 1.38 \times 10^{-3}$ J/K is the Boltzmann constant and T is the temperature [20]. Cantilevers with different shape require finite element analysis, although in few cases analytical formulas can be derived. Neumeister and Ducker [21] obtained approximate analytical expressions for the normal and the lateral spring constant of V-shaped cantilevers. Surfaces with well-defined profiles are useful *in-situ* tools for calibration [22]. In such a case the the horizontal and vertical components of the total force revealed by the photodetector differ from the normal and lateral components with respect to the surface, F_N and F_L. The geometric relations between these components are used to determine the conversion ratio between volts and nanonewtons as well as the coefficient of friction.

In some cases an adequate estimation of the radius of curvature of the tip, R, is required. This quantity can be evaluated with a scanning electron microscope. Otherwise, well-defined structures such as step sites [23] or whiskers [24] can be imaged. The image of these high-aspect ratio structures is convolved with the tip shape. A deconvolution algorithm that allows extracting the radius of curvature of the probing tip was proposed by Villarrubia [25].

Finally, when calibrating the instrument one must be aware that the accuracy of the piezoelectric scanners is limited by instrumental drifts [26] and typical piezoelectric effects such as nonlinearity, hysteresis, creep, and variation of sensitivity with the applied voltage [27, 28]. An accurate analysis of the errors occurring in the calibration of the lateral forces is given in [29].

2.3. Other Techniques

Beside the atomic force microscope, the *surface force apparatus* (SFA) and the *quartz crystal microbalance* (QCM) are useful instruments in nanotribology.

The surface force apparatus [30, 31] consists of a pair of mica sheets, which are pressed together and reciprocally translated under pressure (Fig. 6). The contact area and distance between mica surfaces can be measured by optical or capacitive techniques with a resolution of 2 Å. The normal and lateral forces are deduced from the deformation of springs. The SFA is commonly used to detect the behavior of lubricant liquids between the two surfaces in contact. As an example, the layering of liquids in discrete strata was observed as a function of the applied load [33]. O'Shea et al. observed a similar effect with an AFM [34, 35]. In such a case, the layering of liquid molecules on graphite and mica surfaces caused stiffness oscillation as the tip–sample distance was varied.

The quartz crystal microbalance, which is commonly used to measure thin film growth, was first applied to nanotribology by Krim and Widom [36]. The QCM consists of a single crystal of quartz, which oscillates in a shear mode with a high quality factor Q. The sliding friction of adsorbed films

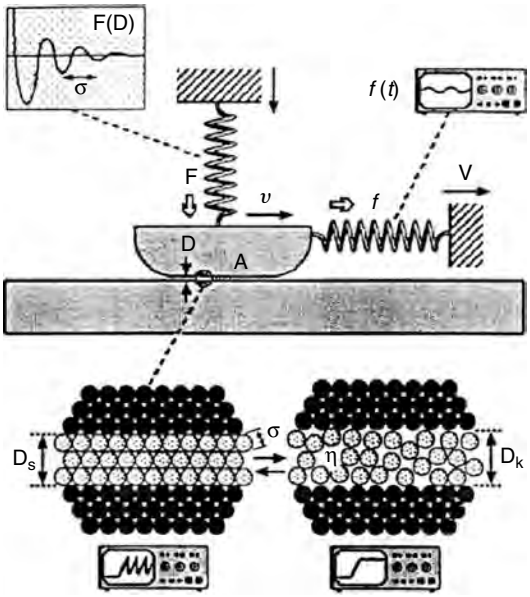


Figure 6. Diagram of a surface force apparatus. Two smooth mica surfaces are separated by a molecularly thin lubricant film. The springs represent any device used to measure the normal and lateral forces between the surfaces. Reprinted with permission from [32], B. Bhushan et al., *Nature* 374, 607 (1995). © 1995, Nature Publishing Group.

is related to the frequency shift and the broadening of the resonance peak of the crystal. Unfortunately, the instrument can reveal only very low friction forces and its application is limited to systems like rare-gas solids on metals [37]. In Section 6.5 we will see how the quartz crystal microbalance was used to detect electronic contributions to friction.

3. MECHANICS OF SLIDING NANOCONTACTS

3.1. Load Dependence of Friction

The lateral force, F_L, between two surfaces in reciprocal motion depends on the size of the *real* area of contact, A, which is usually a few orders of magnitude smaller than the apparent area of contact. From the analysis of several experimental data Bowden and Tabor concluded that friction is proportional to the real area A [38]:

$$F_L = \sigma A \tag{5}$$

The mean lateral force per unit area, σ, is called the *shear strength*. In a first approximation, it does not depend on the normal pressure. Bowden and Tabor assumed that the asperities of the two surfaces in contact are deformed *plastically* while sliding. In such a case the asperities of the two surfaces are compressed until the applied pressure, p, equals a certain yield value, p^*. The resulting area of contact is thus $A = F_N/p^*$, and the well-known Amonton's law is obtained: $F_L = \mu F_N$, where $\mu = \sigma/p^*$ is the *coefficient of friction*. The same idea can be extended to contacts formed by many asperities, and it leads again to Amonton's law. The simplicity of this analysis explains why most of the friction processes

were related to plastic deformation for a long time. However, plastic deformation provokes a quick damage of the surfaces under pressure, which is not so frequently observed while sliding. Thus, the morphology of the surface must be affected by less disruptive mechanisms of deformation.

Elastic deformation can be easily described in the case of a sphere of radius R pressed against a flat surface. In such a case the contact area is [39]

$$A(F_N) = \pi\left(\frac{R}{K}\right)^{2/3} F_N^{2/3} \tag{6}$$

where $K = 3E^*/4$, E^* being an effective Young's modulus, related to the Young's moduli, E_1 and E_2, and the Poisson's numbers, ν_1 and ν_2, of the sphere and plane, by the following relation:

$$\frac{1}{E^*} = \frac{1-\nu_1^2}{E_1} + \frac{1-\nu_2^2}{E_2} \tag{7}$$

The result $A \propto F_N^{2/3}$ does not agree with Amonton's law, $A \propto F_N$. However, a linear relation between F_L and F_N can be derived for contacts formed by several asperities in some important cases. For example, the area of contact between a flat surface and a set of asperities with an exponential height distribution and the same radius of curvature, R, is proportional to the normal force [5]. The same conclusion holds approximately for a Gaussian height distribution. However, the hypothesis that the radii of curvature are the same for all asperities is not realistic. A general model proposed by Persson predicts that the contact area is proportional to the applied load for a large variety of elastoplastic contacts formed by surfaces with arbitrary roughness [6].

Further effects are observed if *adhesion* between the asperities is taken into account. If the range of action of the adhesive forces is smaller than the elastic deformation, Eq. (6) is extended by the *Johnson–Kendall–Roberts* (JKR) *model* [40] to the expression

$$A(F_N) = \pi\left(\frac{R}{K}\right)^{2/3}\left(F_N + 3\pi\gamma R + \sqrt{6\pi\gamma R F_N + (3\pi\gamma R)^2}\right)^{2/3} \tag{8}$$

where γ is the surface tension between sphere and plane. The real area of contact at zero load is finite and the sphere can be detached only by pulling it away with a finite force. This is also true in the opposite case, where the range of action of adhesive forces is larger than the elastic deformation. In such a case the relation between contact area and load has the simple form

$$A(F_N) = \pi\left(\frac{R}{K}\right)^{2/3}\left(F_N - F_{\text{off}}\right)^{2/3} \tag{9}$$

where F_{off} is the negative load that is required to break the contact. The Hertz-plus-offset relation (9) can be derived from the *Derjaguin–Muller–Toporov* (DMT) *model* [41]. To discriminate between the JKR or DMT models Tabor [42] introduced a non-dimensional parameter,

$$\Phi = \left(\frac{9R\gamma^2}{4K^2 z_0^3}\right)^{1/3} \tag{10}$$

where z_0 is the equilibrium distance in contact. The JKR model can be applied if $\Phi > 5$; the DMT model holds when $\Phi < 0.1$. For intermediate values of Φ, the *Maugis–Dugdale model* [43] is consistent with experimental results.

The tip of a friction force microscope is a single asperity sliding on a surface. The previous discussion suggests a nonlinear dependence of friction on the applied load, provided that continuum mechanics can be applied down to nanometer scales. Schwarz et al. [44] observed that the Hertz-plus-offset relation (9) holds on graphite, diamond, amorphous carbon, and C_{60} in argon atmosphere (Fig. 7). Well-defined spherical tips with radii of curvature of tens of nanometers were used. The probing tips were obtained by contaminating standard silicon tips with amorphous carbon in a transmission electron microscope. In order to compare the tribological behavior of different materials, Schwarz et al. introduced an *effective coefficient of friction*, $\widetilde{C}$, which does not depend on the tip curvature. The highest friction coefficient $\widetilde{C}$ was found on C_{60} and the lowest on graphite.

Various groups measured friction vs load curves in UHV in agreement with JKR theory [45–47]. Different materials such as ionic crystals, mica, and metals were considered. Meyer et al. used an original two-dimensional (2D) histogram technique [45]. The normal load is increased or decreased while imaging, which allows one to correlate lateral and normal forces with improved statistics. Carpick et al. extended the JKR relation (8) to include nonspherical tips. For example, in the case of an axisymmetric tip profile $z \propto x^{2n}$, with $n > 1$, the increase of friction becomes less pronounced with increasing n [46].

3.2. Estimation of the Contact Area

In contrast to other tribological instruments, like the surface force apparatus, the area of contact cannot be directly determined with the friction force microscope. An indirect estimation of the contact area is achieved by *contact stiffness* measurements [48]. Figure 8 shows a series of two springs corresponding to the microscope tip apex pressed

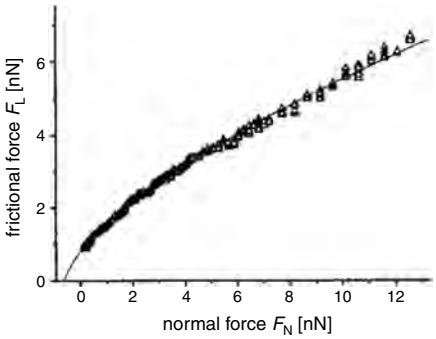


Figure 7. Friction vs load curve measured on amorphous carbon with a well-defined spherical tip in argon atmosphere. The data are fitted with the expression (9). Reprinted with permission from [44], U. D. Schwarz et al., *Phys. Rev. B* 56, 6987 (1997). © 1997, American Physical Society.

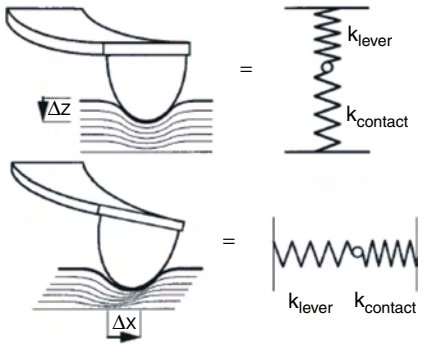


Figure 8. (a) Normal and (b) lateral stiffness of the contact between tip and surface. Reprinted with permission from [49], R. W. Carpick et al., *Appl. Phys. Lett.* 70, 1548 (1997). © 1997, American Institute of Physics.

against the sample. The effective constant k^z_{eff} of the series is given by

$$\frac{1}{k^z_{\text{eff}}} = \frac{1}{k^z_{\text{contact}}} + \frac{1}{c_N} \tag{11}$$

where c_N is the normal spring constant of the cantilever, and k^z_{contact} is the contact stiffness. This quantity is related to the radius of the contact area, a, by the simple relation

$$k^z_{\text{contact}} = 2aE^* \tag{12}$$

where E^* is the effective Young's modulus introduced in Section 3.1. Unfortunately, typical values of k^z_{contact} are order of magnitudes larger than c_N, and a practical application of Eq. (11) is not possible.

Carpick et al. [49] and Lantz et al. [50, 51] extended this idea to *lateral* deformations. According to several models, the lateral stiffness of the contact between a sphere and a flat surface is [52]

$$k^x_{\text{contact}} = 8aG^* \tag{13}$$

where the effective shear modulus G^* is defined by

$$\frac{1}{G^*} = \frac{2-\nu_1^2}{G_1} + \frac{2-\nu_2^2}{G_2} \tag{14}$$

G_1, G_2 are the shear moduli of sphere and plane. The contact between tip and sample can be modeled again by a series of springs (Fig. 8b). The effective constant k^x_{eff} of the series is now given by

$$\frac{1}{k^x_{\text{eff}}} = \frac{1}{k^x_{\text{contact}}} + \frac{1}{k^x_{\text{tip}}} + \frac{1}{c_L} \tag{15}$$

where c_L is the lateral spring constant of the cantilever, and k^x_{contact} is the lateral stiffness of the contact. As suggested by Lantz et al., Eq. (15) includes the lateral stiffness of the tip, k^x_{tip}, which can be of the same order of the lateral spring constant. The effective spring constant k^x_{eff} is determined by scanning the surface and measuring the slope of the lateral force vs displacement curve (Section 5.1). Once k^x_{contact} is known, the contact radius, a, is easily estimated with Eq. (13).

Carpick et al. applied their method to a silicon nitride tip sliding on muscovite mica in air, whereas Lantz et al. studied the contact between $NbSe_2$ and graphite in UHV. The relations between spring constant and load were fitted by the JKR and the Maugis–Dugdale models respectively. The same models were also consistent with independent measurements of friction vs load. Thus, the hypothesis that friction is proportional to the contact area held in the applied range of loads (up to $F_N = 40$ nN in both experiments).

Another way to estimate the contact area is based on the measurement of the contact conductance, which is proportional to the contact area. Enachescu et al. used silicon cantilevers coated with tungsten carbide to measure both conductance and friction on boron-doped diamond as a function of the applied load [53, 54]. The load dependences of both friction and conductance were fitted independently with the same model (DMT), which was a further confirmation of the hypothesis (5).

4. FRICTION EXPERIMENTS ON THE NANOMETER SCALE

Friction force microscopy is mainly applied to characterize the response of a surface to the localized stress exerted by the probing tip. The large variety of materials investigated by FFM includes solid lubricants, carbon and silicon compounds, ferroelectrics, and even ice. The role of the environment in such a contest is essential. Ambient air, controlled atmosphere or liquids [55], can drastically influence the values of friction forces, as well as their dependence on the scan load and velocity. Only ultra-high-vacuum friction results from the direct interaction of two materials in contact. Atomic-scale experiments in UHV will be discussed later in Section 6. In the same section we will also consider experiments dealing with friction anisotropy and other minor effects.

4.1. Material Contrast Revealed by Friction Force Microscopy

Due to their lubricity properties, *self-assembled monolayers* (SAM) and, in particular, *Langmuir–Blodgett* (LB) *films* are among the most studied materials in nanotribology. LB films can be easily prepared by transferring any desired number of layers on smooth substrates, as silicon wafers or mica, where they form densely packed structures. Figure 9 shows the first FFM measurements on LB films realized by Meyer et al. [56]. Two bilayers of Cd–arachidate were transferred on a silicon wafer. Friction is lower on the areas covered by the lubricant film, and it does not depend on the film thickness. The load dependence of friction is weak, and wear appears when the normal load exceeds 10 nN. Furthermore, small islands of bilayer height could be moved in their entirety, which made possible a direct measurement of the shear strength between the bilayers. The measured value of $\sigma =$ 1 MPa was in agreement with SFA experiments at low loads [57]. Overney et al. [58] studied more complex LB films, formed by mixtures of hydrocarbons and fluorocarbons. Fluorocarbons were observed to be less effective in reducing friction than hydrocarbons, which was again in agreement

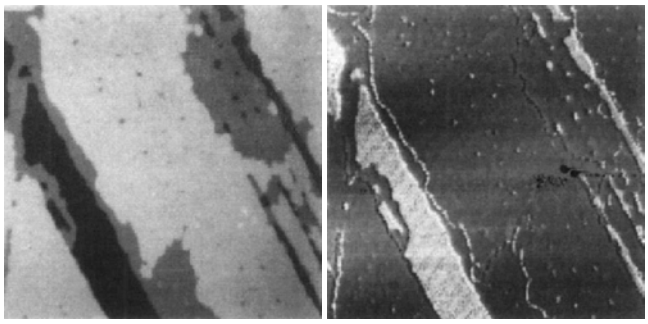


Figure 9. (a) Topography of two bilayers of Cd–arachidate on a silicon wafer. (b) Friction map shows increased friction on the substrate compared to the substrate. Friction is the same on the single and on the couble bilayer. Frame sizes: 2 μm. Reprinted with permission from [56], E. Meyer et al., *Phys. Rev. Lett.* 69, 1777 (1992). © 1992, American Physical Society.

with SFA measurements. Frisbie et al. [59] studied the relation between friction and chemistry on SAMs and found different contrasts when the tip was coated with different functional groups. An extended study with the same purpose was performed by Green et al. [60]. A detailed overview on this topic was given by Carpick and Salmeron [61].

Carbon surfaces were first observed by Mate [62], who studied friction on C_{60}, amorphous carbon, hydrogen terminated diamond, and graphite. The highest friction was found on C_{60}, and the lowest on graphite, which is in agreement with the results by Schwarz et al. (Section 3.1). Buzio et al. [63] applied the idea of the effective friction coefficient introduced by Schwarz et al. to characterize friction on nanostructured carbon films with self-affine morphology. Perry et al. [64] observed that friction on amorphous carbon films increases with increasing hydrogen content. These results are of great interest in the hard disk drive industry, where carbon protective coatings are applied. Lüthi et al. [65] investigated friction on C_{60} islands deposited on NaCl in UHV. Again C_{60} revealed high friction. Furthermore, the islands could be moved over the substrate without breaking and with extraordinary low shear strength $\sigma =$ 0.05–0.1 MPa (Fig. 10). This result suggested to the authors

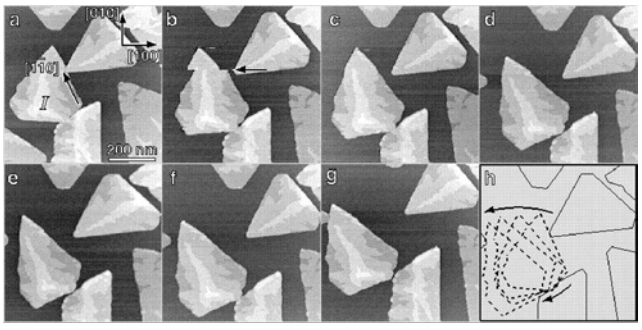


Figure 10. (a)–(g) Sequence of topography images of C_{60} islands on NaCl. An island is rotated and translated from one location to another. Frame sizes: 530 nm. (h) Summary of the movement of the C_{60} island. Reprinted with permission from [65], R. Lüthi et al., *Science* 266, 1979 (1994). © 1994, American Association for the Advancement of Science.

that C_{60} islands might be used as a sled-type transport system on the nanometer scale.

Silicon and silicon oxides have a fundamental role in semiconductor industry. Apart from their electronic properties, mechanical properties are interesting in the development of microdevices, where low friction is required to minimize power consumption. Scandella et al. [66] examined a silicon surface structured by standard photomask lithography and found that friction on hydrogen passivated silicon was larger by a factor 2 than friction on silicon oxide. Teuschler et al. [67] patterned silicon surfaces with a conductive tip, which was subsequently used for characterizing friction. Friction was higher on the modified areas, where the formation of silicon oxide was enhanced. The apparent contrast with the experiment of Scandella et al. is probably due to different crystallinity of the samples. Studies on III–V semiconductors were reported by Garcia and co-workers. Tamayo et al. [68] detected chemical variations of indium alloys with 3 nm resolution. Changes of 10% in indium composition were clearly distinguished in air. Measurements with submonolayer sensitivity were performed on quantum dot structures [69].

Friction force microscopy was also applied to investigate ferroelectric materials. Lüthi et al. [70] and Eng et al. [71] observed significant contrasts between neighboring domains of opposite polarization in friction force maps acquired on GASH (guanidimium aluminum sulphate hexahydrate) and TGS (triglycine sulphate), respectively. On TGS Bluhm et al. measured different friction coefficients, depending on the polarization, the asymmetry of the surface potentials, and also the orientation of the crystallographic lattice with respect to the scan direction [72]. On GASH, the contrast was related only to structural differences, which modify the surface potential experienced by the probing tip. The electrostatic interaction between tip and sample did not affect the friction force significantly.

To conclude this overview, we should also mention the studies of friction on a nanometer thin ice film grown on mica by Bluhm et al. [73]. A friction coefficient $\mu = 0.60$ was measured in a temperature range from -24 to -40 °C, which is comparable to the values obtained in macroscopic experiments [38]. The squeezing of the water layer out of the contact, which was observed in noncontact mode, suggested that dry friction was probably revealed. Other effects like pressure melting and frictional heating were found to be not relevant.

4.2. Role of the Environment

A humidity increase in ambient air leads to the formation of water layers on hydrophilic surfaces and to a consequent capillary interaction between tip and surface [74]. The influence of capillary condensation in FFM measurements was first studied by Binggeli and Mate [75], who found a substantial decrease of friction at 70% humidity on a hydrophilic silicon oxide surface. On less hydrophilic amorphous carbon films and lubricated silicon oxide, the decrease was not observed. Systematic studies of humidity dependence on mica were performed by Schumacher et al. [76]. The formation of an ordered layer of adsorbed water at 40% humidity led to a strong increase in friction. The Hertzian

dependence $F_L \propto F_N^{2/3}$ was also observed. However, friction decreased above 60%, probably because water acted as a boundary lubricant. Putman et al. [77] found out the relation $F_L \propto F_N^{2/3}$ on mica and glass in ambient conditions, whereas friction increased linearly with the normal load in N_2 or Ar atmosphere. This result was explained assuming that the tip roughness was smoothed by a condensed water film, which led to a single asperity contact at high humidity and to a multiasperity contact at low humidity. Capillary condensation leads also to different velocity dependences of friction on the nanometer scale. This effect is discussed in Section 6.3. We should also mention that friction in aqueous solutions depends on the pH [78, 79]. Thus, in principle, it might be possible to differentiate chemical species on a surface simply by changing the pH value of the surrounding medium.

4.3. Topographic Effects

Another factor to take into account when measuring friction with an atomic force microscope is the contribution of the topography to the lateral force signal, which in such a case is not simply proportional to the friction force acting on the probing tip. We have already seen how a surface with well-known local slope the topographic contribution can be used to calibrate the lateral forces (Section 2.2). This is not the case with rough surfaces, where the slope variations, combined with local variations of the contact area, can lead to ambiguous results. These problems have been addressed by Bhushan and co-workers [80, 81].

A particular topography effect is observed at the step edges of crystal surfaces. Meyer et al. [56] found a significant increase of friction at such locations on a NaCl surface in UHV. The same increase was observed when a step was crossed up- or downward (Fig. 11). This result states that the most important role is not taken by simple geometric effects, but by enhanced interaction at the step locations (*Schwöbel barriers*).

Figure 11. Friction image of NaCl. Increased friction is observed on the step sites, both going down and up the steps, which indicates that friction at the step edges is not dominated by geometric effects. Reprinted with permission from [56], E. Meyer et al., *J. Vac. Sci. Technol. B* 14, 1285 (1996). © 1996, American Vacuum Society.

5. FRICTION ON THE ATOMIC SCALE

Friction forces on the atomic scale are characterized by a typical sawtooth behavior, which reflects the structure of the underlying lattice. This phenomenon is due to the stick-slip movement of the probing tip, and it can be well described by a classical model developed by Tomlinson in 1929 [10]. Theoretical studies of stick-slip on the atomic scale were reported by various groups [82–91]. Here, we first derive the main results of the Tomlinson model in one and two dimensions. Thermal effects, which slightly affect these results at room temperature, will be discussed in Section 5.2.

5.1. The Tomlinson Model

The motion of the AFM tip is ultimately related to the atomic structure of the surface and the elastic deformations of the cantilever. The shape of the tip-surface interaction potential, $U(x, y)$, depends on several factors like the chemical composition of the two materials and the atomic arrangement on the tip apex. For sake of simplicity, we will start the analysis in the one-dimensional case and consider a sinusoidal profile with the periodicity of the atomic lattice, a, and a peak-to-peak amplitude E_0. In Section 3.2, we have shown how the elasticity of the cantilever and of the contact area can be described in an unique framework introducing the effective lateral spring constant k_{eff}. If the cantilever moves with a constant velocity V along the x direction, the total energy of the system is then given by

$$E_{\text{tot}}(x, t) = -\frac{E_0}{2}\cos\frac{2\pi x}{a} + \frac{1}{2}k_{\text{eff}}(Vt - x)^2 \qquad (16)$$

Figure 12 shows the energy profile $E_{\text{tot}}(x, t)$ at two different instants. When $t = 0$ the tip is localized in the absolute minimum of E_{tot}. This minimum slightly increases with time due to the cantilever motion, until the tip position becomes unstable at a certain time $t = t^*$. The critical position x^* at $t = t^*$ is determined by equating to zero the second derivative $\partial^2 E_{\text{tot}}(x, t)/\partial x^2$, which gives

$$x^* = \frac{a}{4}\arccos\left(-\frac{1}{\gamma}\right) \qquad \gamma = \frac{2\pi^2 E_0}{k_{\text{eff}} a^2} \qquad (17)$$

The coefficient γ compares the strength of interaction between tip and surface with the stiffness of the system. The

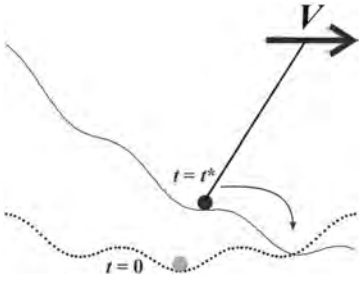


Figure 12. Energy profile experienced by the AFM tip (black circle) at $t = 0$ (dotted line) and $t = t^*$ (continuous line) The tip is dragged by the cantilever, which moves rightward with a constant velocity V. The tip does not move significantly until it suddenly "jumps" at the critical time $t = t^*$.

stick-slip is observed only if $\gamma > 1$ (i.e., if the system is not too stiff or the interaction is strong enough). In what follows, we assume that $\gamma \gg 1$, as usually observed in the experiments. When $t = t^*$ the tip suddenly "jumps" into the next minimum of the potential profile. Just before jumping, the lateral force is [92]

$$F^* = \frac{k_{\text{eff}} a}{2\pi} \tag{18}$$

Figure 13 shows the lateral force F_L as a function of the cantilever position, X. When the cantilever is moved forward, the lower part of the curve in Figure 13 is obtained. If, at a certain point, the direction of motion of the cantilever is suddenly inverted, the force follows the profile in the upper part of the curve. The area of the *friction loop* obtained by scanning back and forth gives the total energy dissipated. The slope of the *sticking* part of the loop is equal to k_{eff} [92].

In two dimensions, the energy of the system is given by

$$E_{\text{tot}}(\mathbf{r}, t) = U(\mathbf{r}) + \frac{k_{\text{eff}}}{2}(\mathbf{V}t - \mathbf{r})^2 \tag{19}$$

where $\mathbf{r} = (x, y)$ and $\mathbf{V}$ is the scan velocity. For example, Figure 14 shows the total energy corresponding to a periodic potential of the form

$$U(x,y) = -\frac{E_0}{2}\left(\cos\frac{2\pi x}{a} + \cos\frac{2\pi y}{a}\right) + E_1 \cos\frac{2\pi x}{a}\cos\frac{2\pi y}{a} \tag{20}$$

The stability of the equilibrium can be discussed introducing the Hessian matrix

$$H = \begin{pmatrix} \dfrac{\partial^2 U}{\partial x^2} + k_{\text{eff}} & \dfrac{\partial^2 U}{\partial x\,\partial y} \\ \dfrac{\partial^2 U}{\partial y \partial x} & \dfrac{\partial^2 U}{\partial y^2} + k_{\text{eff}} \end{pmatrix} \tag{21}$$

When both eigenvalues $\lambda_{1,2}$ of the Hessian are positive the position is stable. Figure 15 shows such regions for a potential of the form (20). The tip follows the cantilever adiabatically as long as it remains in the (++)-region. When the tip is dragged to the border of the region, it suddenly jumps into the next (++)-region.

Thus far we have implicitly assumed that the tip is terminated by only one atom. It is also instructive to consider the

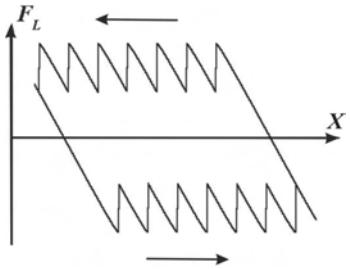


Figure 13. Friction loop obtained by scanning backward and forward. The slope of the sticking part of the loop gives the effective spring constant k_{eff}.

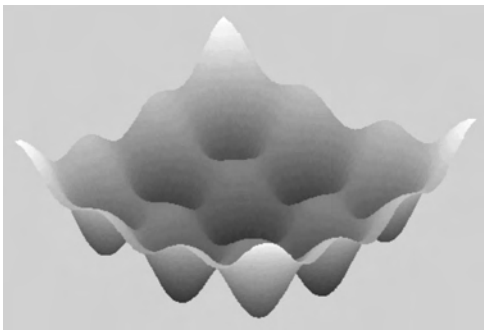


Figure 14. Energy landscape experienced by the AFM tip in 2D.

case of a periodic surface sliding on another periodic surface. In the *Frenkel–Kontorova–Tomlinson* (FKT) *model* the atoms of one surface are harmonically coupled with their nearest neighbors. Let us examine the case of quadratic symmetries, with lattice constants a_1 and a_2 for the upper and lower surface respectively (Fig. 16). In such a context the role of commensurability is essential. In one dimension Weiss and Elmer [94] predicted that friction should decrease with decreasing commensurability, the minimum of friction being reached when a_1/a_2 equals the *golden mean* $\bar{z} = 1 + (1 + (1 + \cdots)^{-1})^{-1} = 0.618$. In two dimensions Gyalog and Thomas [95] studied the case $a_1 = a_2$, with a misalignment between the two lattices given by an angle θ. When the sliding direction changes, friction increases from a minimum value corresponding to the sliding angle $\varphi = \theta/2$ to a maximum value, which is reached when $\varphi = \theta/2 + \pi/4$. The misfit angle θ is related to commensurability. Since the misfit angles giving rise to commensurate structure form a dense subset, the dependence of friction on θ should be discontinuous. The numerical simulations performed by Gyalog agree with these conclusions (Fig. 17).

5.2. Thermal Effects on Atomic Friction

Although the Tomlinson model gives a good interpretation of the basic mechanism of the atomic stick-slip, it cannot explain some minor features observed in the experiments.

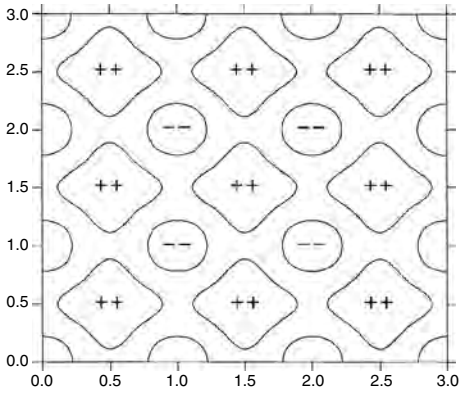


Figure 15. Regions on the (x, y) plane labeled according to the signs of the eigenvalues of the Hessian matrix (21) with the potential (20). Reprinted with permission from [93], T. Gyalog et al., *Europhys. Lett.* 31, 269 (1995). © 1995, EDP Sciences.

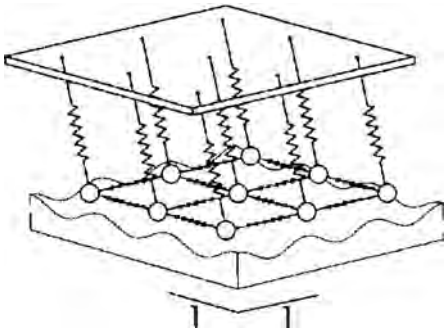


Figure 16. The FKT model in 2D. Reprinted with permission from [95], T. Gyalog and H. Thomas, *Europhys. Lett.* 37, 195 (1997). © 1997, EDP Sciences.

For example, the peaks of the friction loop shown in Figure 2a have different heights, which is not the case in Figure 13. Another effect is observed if the scan velocity V is varied. In Section 6.3 we will see that the mean friction force measured in UHV conditions increases with the logarithm of V (Fig. 18). Also this result cannot be interpreted without further assumptions in the model.

Let us focus again on the energy profile in Figure 12. At any time $t < t^*$, the tip jump is prevented by the energy barrier $\Delta E(t) = E(x_{\max}, t) - E(x_{\min}, t)$, where $x_{\max}(t)$ corresponds to the first maximum observed in the energy profile, and $x_{\min}(t)$ is the actual position of the tip (Fig. 19). The quantity ΔE decreases with time, or, equivalently, with the frictional force F_L, until it vanishes when $F_L = F*$. Close to the critical point the energy barrier can be approximated as

$$\Delta E(F_L) = \lambda(\widetilde{F} - F_L) \tag{22}$$

where $\widetilde{F} \approx F^*$. However, at a finite temperature T, the jump can occur *before* the critical value $F_L = F^*$ is reached [9]. The most probable value of F_L, when the jump occurs, is estimated with the *master equation*

$$\frac{dp(t)}{dt} = -f_0 \exp\left(-\frac{\Delta E(t)}{k_B T}\right) p(t) \tag{23}$$

In (23), p is the probability that the tip does *not* jump and f_0 is a characteristic attempt frequency of the system. Using

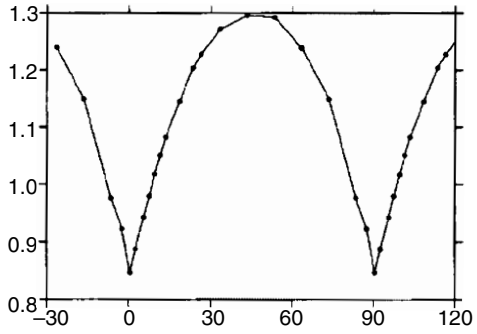


Figure 17. Friction as a function of the sliding angle φ in the 2D FKT model. Reprinted with permission from [95], T. Gyalog and H. Thomas, *Europhys. Lett.* 37, 195 (1997). © 1997, EDP Sciences.

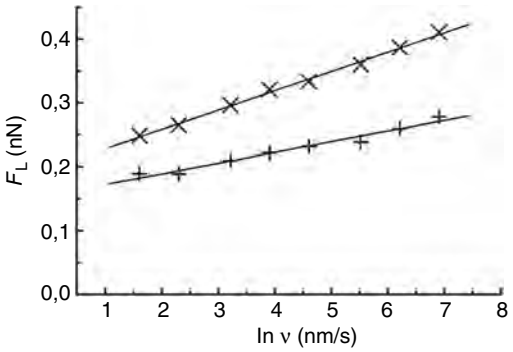


Figure 18. Mean friction force vs scanning velocity on NaCl(100) at $F_N = 0.44$ nN (+) and $F_N = 0.65$ nN (×). Adapted with permission from [9], E. Gnecco et al., *Phys. Rev. Lett.* 84, 1172 (2000). © 2000, American Physical Society.

the approximation (22) the condition of highest probability of jumping, $d^2p(F)/dF_L^2 = 0$, yields [9]

$$F_L(v) = F^* - \frac{k_B T}{\lambda} \ln \frac{V_c}{V} \tag{24}$$

with

$$V_c = \frac{f_0 k_B T}{k_{\text{eff}} \lambda} \tag{25}$$

Thus, the lateral force increases logarithmically with the sliding velocity. However, when the jump occurs very close to the critical point $x = x^*$ the approximation (22) does not hold. This is the case at high velocities, where the probability $p(t)$ does not change significantly, until it suddenly falls down to zero when $t \to t^*$. Thus, friction is expected to be constant at sufficiently high velocities, which agrees with the classical Coulomb's law. This topic is still under investigation.

5.3. Molecular Dynamics Simulations of Friction and Wear

Several authors interpreted sliding friction on the atomic scale with molecular dynamics (MD). Even if huge differences in time-scale make impossible a direct comparison between simulations and real experiments, MD simulations can reveal atomic mechanisms not directly accessible by experiment. Landman et al. [96, 97] predicted atomic-scale

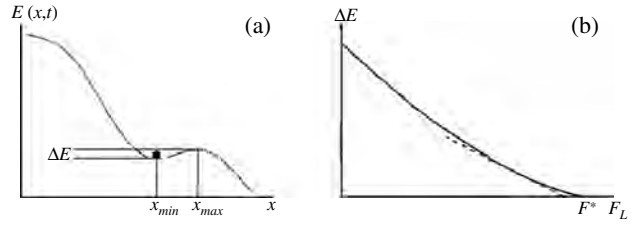


Figure 19. (a) The tip jump in the Tomlinson model at any time $t < t^*$ is prevented by an energy barrier $\Delta E(t)$. (b) Energy barrier ΔE as a function of the lateral force F_L. The dashed line close to the critical value corresponds to the linear approximation (22).

stick-slip for Si and CaF_2 tips sliding on surfaces of the same materials. Harrison et al. [98] observed stick-slip with a weak load dependence for two hydrogen-terminated diamond (111) surfaces, which is in a certain agreement with the experiments by Germann et al. discussed in Section 6.1. Sørensen et al. [99] predicted the occurrence of wear in the case of a copper tip sliding on Cu(100) surfaces (Fig. 20), whereas Cu(111) appeared more resistant. Even this conclusion was somehow confirmed experimentally (Section 6.1). Shluger et al. observed that the scanning process is accompanied by strong displacements of the surface ions inside the lattice and by their transient or permanent adsorption onto the tip at low loads (till 1 nN) [100–102]. Another simulation by Livshitz and Shluger [103] suggested that the adsorbed material adjusts itself leading to an effect of self-lubrication of the tip (Fig. 21). This result may explain why periodic structures can be observed experimentally also at rather high loads. Ohzono and Fujihira [104, 105] considered friction between an ordered organic monolayer of $n-$alkane molecules and a rigid slider with a single protuberance. Incommensurability and tip size comparable to the molecular size were found to be important conditions in imaging the molecular lattice.

A few MD simulations focused on the connection between friction and wear on the nanometer scale. Buldum and Ciraci [106] considered the processes of nanoindentation and sliding of sharp and blunt nickel tips on copper. In the case of the sharp tip quasi-periodic variations of the lateral force were observed, due to stick-slip involving phase transition. One layer of the asperity was deformed to match the substrate during the first slip and then two asperity layers merged into one through structural transition during the second slip. In the case of the blunt tip the stick-slip was less regular. Different mechanisms were observed when the tip is harder than the underlying sample. Komanduri et al. [107] considered an infinitely hard Ni indenter scratching single crystal aluminum at extremely low depths (Fig. 22). In such case a linear relation between friction and load was found, with a high coefficient of friction $\mu = 0.6$, independent of the scratch depth. Nanolithography simulations were performed by Fang et al. [108], who investigated the role of the displacement of the probing tip along the direction of slow

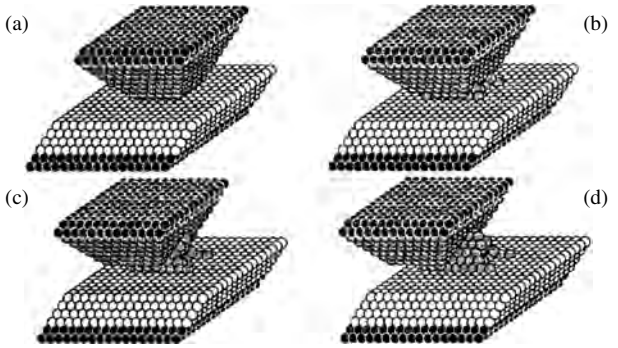


Figure 20. Snapshot of a Cu(100) tip on a Cu(100) substrate during sliding. (a) Starting configuration; (b,c,d) snapshots after two, four, and six slips. Reprinted with permission from [99], M. R. Sørensen et al., *Phys. Rev. B* 53, 2101 (1996). © 1996, American Physical Society.

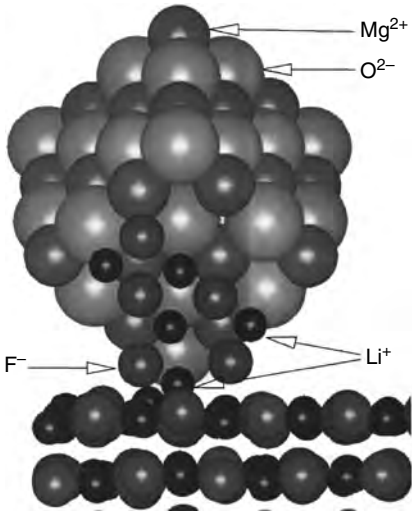


Figure 21. Regular structure of ions adsorbed on a MgO tip sliding on a LiF surface. Reprinted with permission from [103], A. I. Livshits et al., *Phys. Rev. B* 56, 12482 (1997). © 1997, American Physical Society.

motion between a scan line and the next one. A certain correlation with FFM experiments on silicon films coated with aluminum was found. More complex simulations of nanoindentation are presented in [109–112].

6. FRICTION EXPERIMENTS ON THE ATOMIC SCALE

6.1. Experimental Overview

Figure 23 shows the friction map obtained by Mate et al. on graphite on the atomic scale [8]. The periodicity of the lateral force is the same as of the atomic lattice of the sample. The series of friction loops in Figure 24 reveal the stick-slip effect discussed in Section 5.1 The applied loads are in the range of tens of μN. According to the continuum models introduced in Section 3 these values correspond to contact diameters of 100 nm. A possible explanation for the atomic

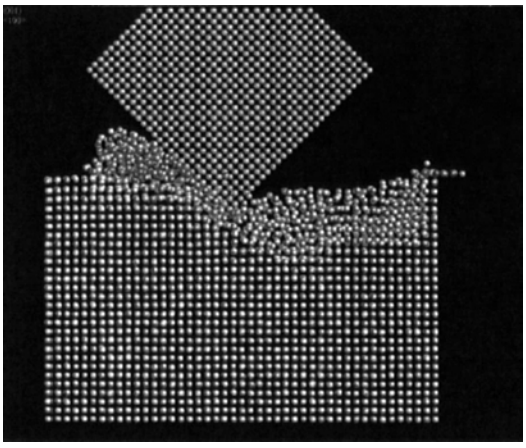


Figure 22. MD simulation of a scratch realized with an infinitely hard tool. Reprinted with permission from [107], R. Komanduri et al., *Phys. Rev. B* 61, 14007 (2000). © 2000, American Physical Society.

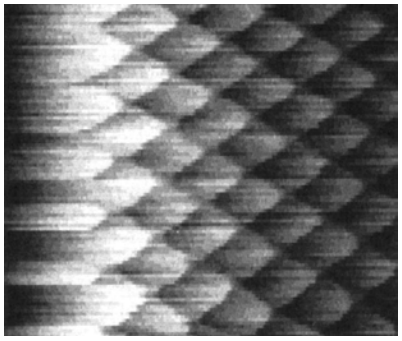


Figure 23. First atomic friction map acquired on graphite with a normal force $F_N = 56$ μN. Frame size: 2 nm. Reprinted with permission from [8], C. M. Mate et al., *Phys. Rev. Lett.* 59, 1942 (1987). © 1987, American Physical Society.

features observed at such high loads is that graphite flakes were detached from the surface and adhered to the tip apex [113]. Another possibility is that the contact between tip and surface consisted of few nm-scale asperities and the corrugation was not entirely averaged out while sliding. The load dependence of friction found by Mate et al. was rather linear, with a small friction coefficient $\mu = 0.01$.

The UHV environment reduces the influence of contaminants on the surface and leads to more precise and reproducible measurements. Howald et al. obtained a series of results on ionic crystals with a home-built AFM apparatus in UHV [114]. In Figure 25 a friction map recorded on KBr(100) is compared with a theoretical map deduced from the 2D-Tomlinson model [115]. The periodicity $a = 0.47$ nm corresponds to the spacing between equally charged ions. No

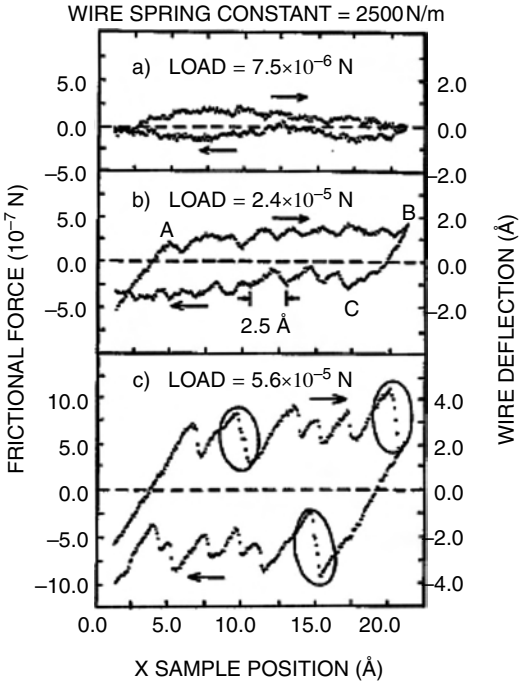


Figure 24. Friction loops on graphite acquired with (a) $F_N = 7.5$, (b) 24, and (c) 75 μN. Reprinted with permission from [8], C. M. Mate et al., *Phys. Rev. Lett.* 59, 1942 (1987). © 1987, American Physical Society.

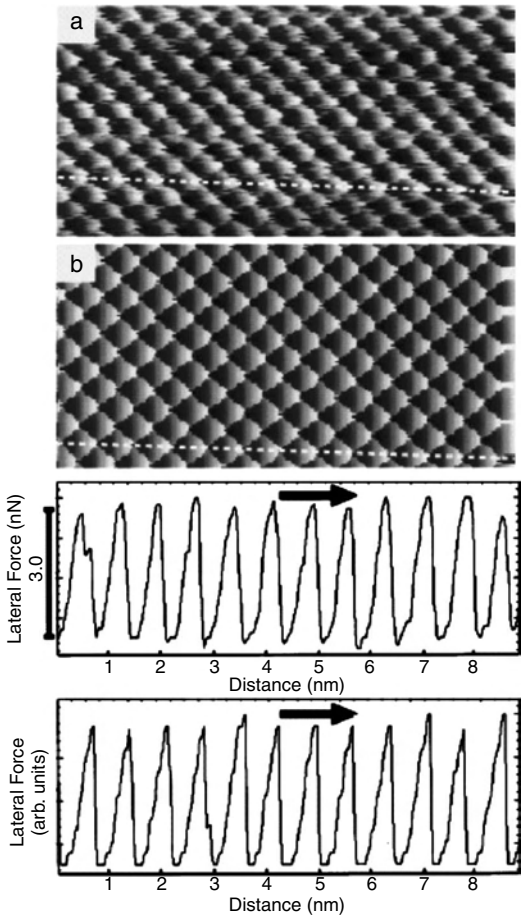


Figure 25. (a) Measured and (b) theoretical friction map on KBr(100). Reprinted with permission from [115], R. Lüthi et al., *J. Vac. Sci. Technol. B* 14, 1280 (1996). © 1996, American Vacuum Society.

individual defects were observed. This is a common aspect to all friction measurements on the atomic scale. One possible reason is that the contact realized by the probing tip is formed by many atoms, which superimpose and average their effects. Molecular dynamics simulations show that even a single-atom contact may cause rather large stresses in the sample, which lead to the motion of defects far away from the contact area. In a picturesque frame, Meyer et al. said that "defects behave like dolphins that swim away in front of an ocean cruiser" [116]. Howald et al. [117] detected atomic-scale friction on the reconstructed Si(111)7×7 surface. The observation of atomic features was only achieved after coating the tips with polytetrafluoroethylene, which has lubricant properties and does not react with the dangling bonds of Si(111)7×7 (Fig. 26). Bennewitz et al. [118] observed atomic stick-slip on metallic surfaces. Regular features were obtained on the closed-packed Cu(111), whereas sliding on Cu(100) produced irregular and irreproducible patterns (Fig. 27). Molecular dynamics suggested that wear occurs more easily on the Cu(100) [110]. This conclusion was achieved with copper tips in computer simulations. The assumption that copper covered the tip apex used in the experiment was supported by current measurements. Atomic stick-slip on diamond was observed by Germann et al. [120]

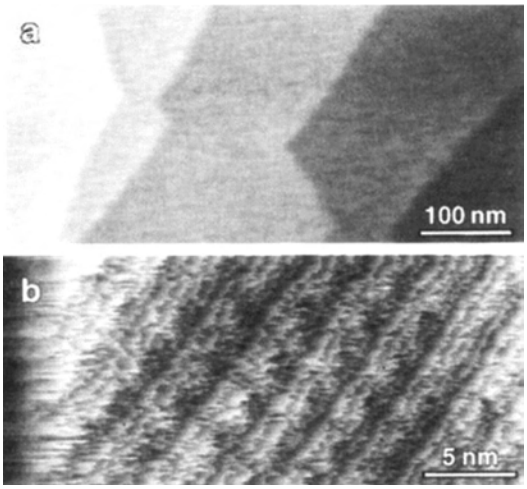


Figure 26. Topography and (b) friction image of Si(111)7 × 7 measured with a PTFE coated Si tip. Reprinted with permission from [117], L. Howald et al., *Phys. Rev. B* 51, 5484 (1995). © 1995, American Physical Society.

with an apposite diamond tip, prepared by chemical vapor deposition, and, a few years later, by van der Oetelaar and Flipse [121] with standard silicon tips. The values of friction had huge discrepancies corresponding to the presence or absence of hydrogen on the surface. Morita et al. [122, 123] measured friction on MoS_2 with a 2D-FFM apparatus, which could reveal forces even perpendicular to the scan direction. The features in Figure 28 correspond to a zigzag walk of the tip, which is consistent with the 2D-Tomlinson model [86, 87]. Analogous studies were conducted on mica and NaF [124]. The 2D stick-slip on NaF was detected with normal forces lower than 14 nN, whereas loads up to 10 μN could be applied on layered materials without damage. The zigzag walk on mica was also observed by Kawakatsu and Saito [125] with an original 2D-FFM with two laser beams and two quadrant photodetectors.

6.2. Friction Anisotropy

The importance of the misfit angle in reciprocal sliding of two flat surfaces discussed in Section 5.1 was first observed by Hirano et al. [126] in the contact between two mica sheets. Friction increased when the two surfaces formed

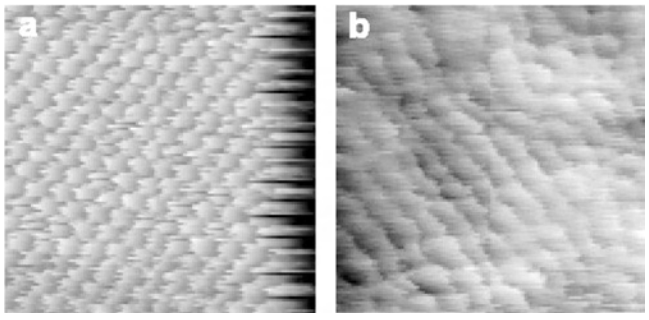


Figure 27. Friction images of (a) Cu(111) and (b) Cu(100). Frame size: 3 nm. Reprinted with permission from [119], R. Bennewitz et al., *Trib. Lett.* 10, 51 (2001). © 2001, Plenum Publishing Corporation.

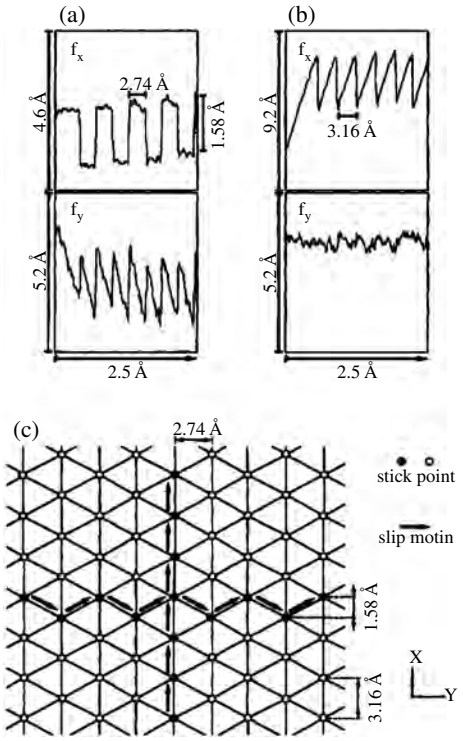


Figure 28. (a) Friction force on MoS_2 acquired by scanning along the cantilever and (b) across the cantilever. (c) Motion of the tip on the sample. Reprinted with permission from [123], S. Fujisawa et al., *Phys. Rev. B* 51, 7849 (1995). © 1995, American Physical Society.

commensurate structures, in agreement with theoretical results. In more recent measurements with a monocrystalline tungsten tip on silicon, Hirano et al. [127] observed *superlubricity* in case of incommensurability. Overney et al. [128] studied the effects of friction anisotropy on a bilayer lipid film and found that different molecular alignments resulted in significant variation of friction. Liley et al. [129] observed islands of a lipid monolayer on mica, consisting of domains with different molecular orientation (Fig. 29). The angular dependence of friction reflected the tilt direction of the alkyl chains of the monolayer revealed with other techniques. In their experiments on C_{60} islands (Section 4.1) Lüthi et al. found that friction is independent of the sliding direction. This was not the case in other experiments by Sheehan and Lieber [130], who observed that the misfit angle is relevant when MoO_3 islands are dragged on the MoS_2 surface. In such a case, sliding was possible only along low index directions. The weak orientation dependence found by Lüthi et al. was probably due to the large mismatch of C_{60} on NaCl. Another example of friction anisotropy is given by carbon nanotubes. A dramatic increase of friction was observed when a nanotube was moved in the directions corresponding to commensurate contact with a graphite surface [131].

6.3. Velocity and Temperature Dependence of Friction

The velocity dependence of friction was investigated relatively late. Zwörner et al. observed that friction between silicon tips and diamond, graphite, or amorphous carbon is

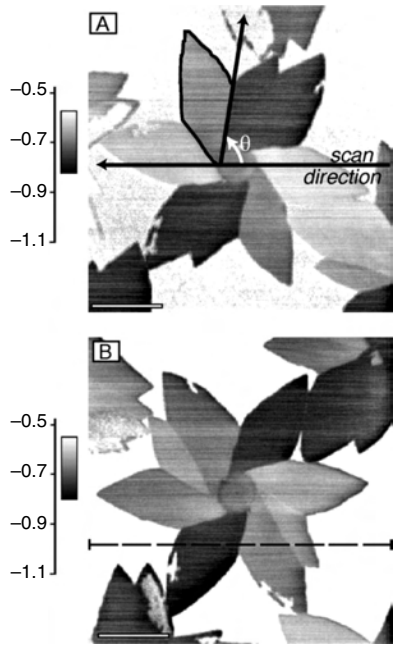


Figure 29. Friction images of a thiolipid monolayer on a mica surface. Reprinted with permission from [129], M. Liley et al., *Science* 280, 273 (1998). © 1998, American Association for the Advancement of Science.

constant with scan velocities, V, of few μm/s [132]. Friction decreased when V was reduced below 1 μm/s. On lipid films on mica, Gourdon et al. [133] found a critical velocity $V_c = 3.5$ μm/s, which discriminated between an increasing friction and a constant friction regime (Fig. 30). Although these results were not explained with thermal activation, we argue that the discussion in Section 5.2 gives the correct interpretative key. A clear observation of a logarithmic dependence of friction on the micrometer scale was reported by Bouhacina et al. [134] on triethoxysilane molecules and polymers grafted on silica with sliding velocity up to $V =$ 300 μm/s The result was explained with a thermally activated

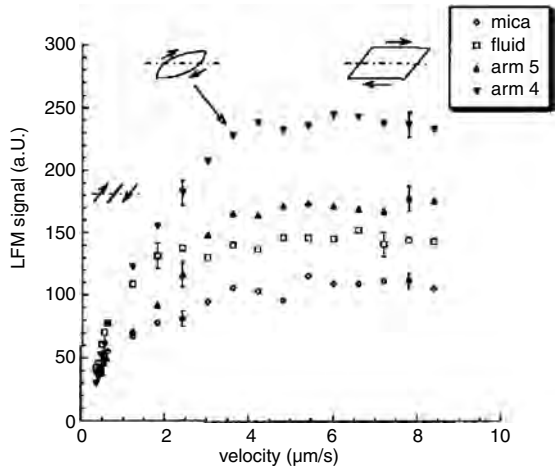


Figure 30. Velocity dependence of friction on mica and on lipid films with different orientation (arms 4 and 5) and in a fluid phase. Reprinted with permission from [133], D. Gourdon et al., *Trib. Lett.* 3, 317 (1997). © 1997, Plenum Publishing Corporation.

Eyring model, which does not differ significantly from the model discussed in Section 5.2 [135, 136]. The first measurements on the atomic scale were performed by Gnecco et al. [9] and Bennewitz et al. [118] on sodium chloride and copper in UHV. In both cases a logarithmic dependence of friction was observed up to $V = 1$ μm/s, in agreement with Eq. (24). Higher values of velocities were not explored, due to the limited range of the scan frequencies accessible with the FFM on the atomic scale. Riedo et al. [137] realized the importance of the surface wettability in the velocity dependence of friction. A logarithmic increase in controlled atmosphere was only observed on hydrophobic surfaces. Hydrophilic surfaces, on the contrary, revealed a logarithmic *decrease* of friction with velocity, which was related to thermally activated nucleation of water bridges between tip and sample asperities. The rate of such a decrease depends drastically on humidity.

Thus far we have used thermal activation to explain the velocity dependence of friction. The same mechanism predicts also that friction should change with temperature. According to the master equation (23) the probability of a tip jump is reduced at low temperatures T until it vanishes when $T = 0$. In this limit case, thermal activation is excluded and the lateral force F_L is equal to F^*, independently of the scanning velocity V. Stick-slip processes at low temperatures have not been reported at our knowledge. However, a significant increase of the mean friction with decreasing temperature was measured by He et al. [138]. Neglecting the logarithmic contributions, Eq. (24) predicts $(F^* - F_L) \sim T$ for the temperature dependence of friction, which is in agreement with the experimental results.

6.4. Dissipation in Noncontact Atomic Force Microscopy

Energy dissipation is also observed when the AFM is operated in noncontact mode (NC-AFM); that is, the cantilever is excited close to its resonance frequency and the energy required to keep the oscillation amplitude constant is monitored [139]. The origin of this additional dissipation is manifold. Apparent energy dissipation may arise from inharmonic cantilever motion, artifacts from the phase controller, or slow fluctuations round the steady-state solution [140, 141]. Velocity dependent dissipation is observed in case of electric- and magnetic-field mediated Joule dissipation [142, 143]. Hysteretic dissipation due to atomic instabilities [144, 145] and hysteresis due to adhesion [146] can also play a role. Energy dissipation can be also measured by exciting the torsional oscillation of the cantilever [147]; in the future it is not excluded that the energy losses observed with FFM and NC-AFM will be considered in a unified frame. A detailed review of dissipation phenomena in noncontact force microscopy is given by [148].

6.5. Electronic Friction

The role of the electrons in friction processes was first suggested by Persson [149]. If small particles of mass m are adsorbed on a thin conducting film the electronic friction

force $F_{L,\mathrm{el}} = m\eta_{\mathrm{el}}v$ gives rise to an additional resistivity

$$\Delta\rho = \eta_{\mathrm{el}}\frac{mn_a}{(ne)^2 d} \tag{26}$$

where n is the number of conduction electrons per unit volume, n_a is the coverage of the adsorbed particles, and d is the film thickness. Experiments in agreement with this effect were performed by Krim et al. [37], who related the tribological properties of thin adsorbed films to changes of the Q-factor of a quartz crystal microbalance. In such way, it was found that slip times for ethane and ethylene monolayers adsorbed on silver and chemisorbed oxygen/silver surfaces are different, arguing that electronic contributions to friction should be considered whenever conducting surfaces are involved [150]. More recently, Dayo et al. [151] proved that friction of incommensurate N_2 on lead drops abruptly at the superconducting transition temperature $T_c \approx 5$ K, which was related again to an electronic contribution to friction. Further experiments are discussed in the review article by Mason [152].

7. WEAR ON THE NANOMETER SCALE

If the normal force F_N applied with the AFM exceeds a critical value depending on the tip shape and the material under investigation, the surface topography is permanently damaged. Among several examples of tip-induced wear of a surface [11, 65, 115, 153–158], wear of the tip was also reported [159, 160]. In some cases wear can be exploited to create patterns with well-defined shape. This is the subject of nanoindentation and nanolithography.

7.1. Nanoindentation

Nanoindentation is performed by increasing and then decreasing the normal load with a hard tip without scanning. The *Berkovich indenter*, formed by a three-sided diamond pyramid, is the most common tool for such a purpose. The indentation pit can be subsequently characterized with the AFM. *Nanohardness* is calculated by dividing the loading force by the projected residual area of the indentation. The geometry of the tip is the most important source of uncertainty in this kind of measurements [161]. Nanohardness tends to decrease when the normal load increases. However, hard and fragile materials react in a different way. For example, Bhushan et al. found that nanohardness falls when $F_N = 35$ mN for Si(111), but not on Si(100) [162].

Besides hardness, the Young's modulus can be obtained from the slope of the unloading curve [163]. A load-displacement curve obtained on Si(100) is shown in Figure 31. The hysteretic appearance means that deformation is not fully elastic. For compliant materials such as polymer blends, magnetic tapes, etc. both loading and unloading curves do not differ.

7.2. Nanolithography

If the AFM tip is scanned with high load, debris can be removed from the underlying surface. Although this technique does not differ from *ploughing* commonly used in agriculture, the results can be surprising when observed on the atomic scale (Fig. 3). Before discussing the few experimental results reported on such a level, we should mention here the importance of nanoscratching in realizing well-defined patterns thanks to the precise control of the tip movement made possible by the piezoelectric elements of the AFM. In such a way, for example, superconducting nanoconstrictions (*Josephson junctions*) [164], or surface quantum wells [165], were created.

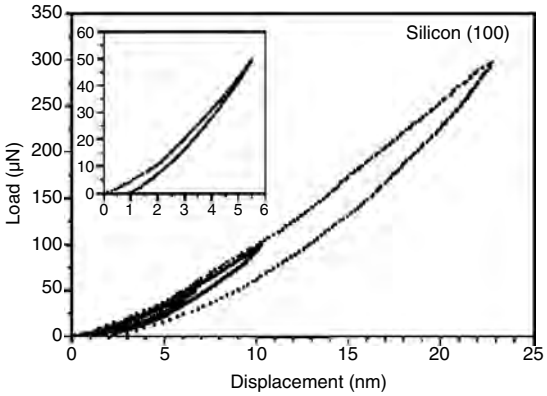


Figure 31. Load-displacement curves at different loads for Si(100). Reprinted with permission from [162], B. Bhushan et al., *Philos. Mag.* 74, 1117 (1996). © 1996, Taylor & Francis, Ltd.

The first atomically resolved images of the damage produced by scratching the AFM tip on an ionic crystal surface were reported by Gnecco et al. [11]. In Figure 3a a small mound grown at the end of a groove on KBr is shown under different magnifications. The groove was created a few minutes before imaging by repeated scanning with the normal force $F_N = 21$ nN. The image shows a lateral force map acquired with a load of about 1 nN; no atomic features were observed in the corresponding topography signal. Figure 3b shows that the debris extracted from the groove recrystallized with the same atomic arrangement of the undamaged surface. Although it is not straightforward to understand how the wear process was initiated and how the tip transported the debris, important indications are provided by the profile of the lateral force F_L recorded while scratching. Figure 32 shows some friction loops acquired when the tip was scanned laterally on 5×5 nm^2 areas. The mean lateral force multiplied by the scanned length gives the total

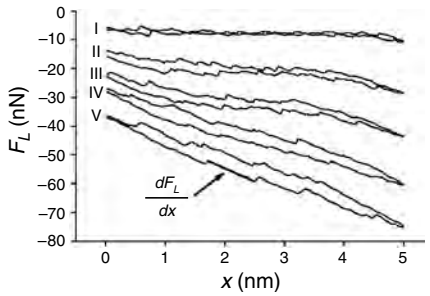


Figure 32. Friction loops acquired while scratching the KBr surface on 5 nm long lines with different loads $F_N = 5.7$ to 22.8 nN. Reprinted with permission from [11], E. Gnecco et al., *Phys. Rev. Lett.* 88, 215501 (2002). © 2002, American Physical Society.

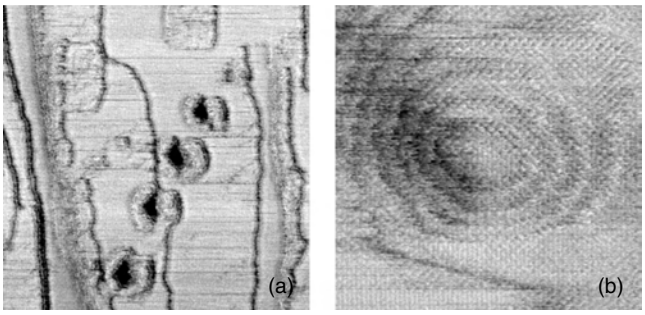


Figure 33. (a) Lateral force images of the pits produced with $F_N = 5.7$ to 22.8 nN. Frame size: 150 nm. (b) Detailed image of one of the pits with pseudo-atomic resolution. Frame size: 20 nm. Adapted with permission from [11], E. Gnecco et al., *Phys. Rev. Lett.* 88, 215501 (2002). © 2000, American Physical Society.

energy dissipated in the process. The results of the tip movement are the pits in Figure 33a. Thanks to the resolution obtained in the lateral force images (Fig. 33b), the number of removed atoms can be directly estimated from the maps. This allowed us to evaluate that 70% of the dissipated energy went into wear-less friction. Figures 32 and 33 clearly show that the damage increased with increasing load. On the other hand, varying the scan velocity between 25 and 100 nm/s did not produce significant changes.

A different kind of wear was observed on layered materials by Kopta and Salmeron [166], who removed layers from a muscovite mica surface with the normal force $F_N =$ 230 nN (Fig. 34a). Fourier filtered images acquired on very small areas revealed the different periodicities of the underlying layers, which reflect the complex structure of the muscovite mica (Fig. 34b, c). To interpret their results, Kopta and Salmeron assumed that wear of mica initiates from atomic defects. When the defects accumulate beyond a critical concentration, they grow to form the scars shown in Figure 34. Such a process was once again related with

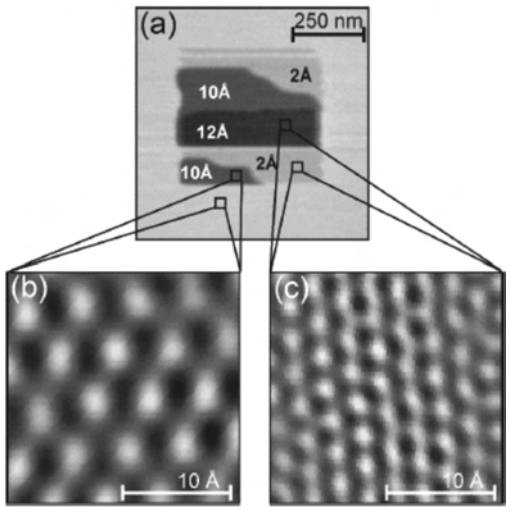


Figure 34. (a) Topography image of an area scratched on muscovite mica with $F_N = 230$ nN. (b,c) Fourier filtered images of different regions. Reprinted with permission from [166], S. Kopta and M. Salmeron, *J. Chem. Phys.* 113, 8249 (2000). © 2000, American Institute of Physics.

thermal activation. The number of defects created in the contact area $A(F_N)$ is [167]

$$N_{\text{def}}(F_N) = t_{\text{res}} n_0 A(F_N) f_0 \exp\left(-\frac{\Delta E}{k_B T}\right) \tag{27}$$

where t_{res} is the residence time of the tip, n_0 is the surface density of atoms, and f_0 is the attempt frequency to overcome the energy barrier ΔE to break a Si–O bond (this quantity depends on the applied load). When the defect density reaches a critical value, a hole is nucleated. The friction force during the creation of a hole was also estimated with thermal activation, which led to the formula [167]

$$F_L = c(F_N - F_{\text{off}})^{2/3} + \gamma F_N^{2/3} \exp(B_0 F_N^{2/3}) \tag{28}$$

The first term on the right gives the wear-less dependence of friction in the Hertz-plus-offset model (Section 3.1); the second term is the contribution of the defect production. The agreement between Eq. (28) and the experiment is clearly visible in Figure 35.

7.3. Tribochemical Wear

Like friction, wear is also influenced by the surrounding environment. Nakahara et al. [168] could induce degradation of the cleavage steps on the hygroscopic crystal $NaNO_3$ at humidity higher than 45%. Diatomic steps splat into monatomic steps and the material was dragged to the terrace above. After repeated scanning, rows of material were piled up. Humidity-dependent layer removal was also observed by Thundat et al. [169] on lead pyrophosphate by scanning at humidity >25%. In such a case the increase of capillary forces at high humidity was proposed as the wear mechanism, instead of the sample dissolution.

7.4. Metallic Nanocontacts

An interesting case of wear processes is provided by nanometer-sized metallic contacts under stress. If the contact is smaller than the mean free path of the electrons, the conductance is given by the Landauer–Buttiker formula [170]

$$G = \frac{2e^2}{h} \sum_k T_k \tag{29}$$

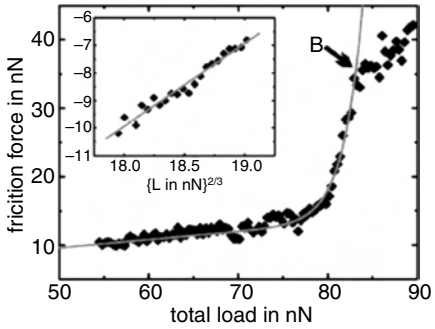


Figure 35. Friction vs load curve during the creation of a hole in the muscovite mica. Reprinted with permission from [166], S. Kopta and M. Salmeron, *J. Chem. Phys.* 113, 8249 (2000). © 2000, American Institute of Physics.

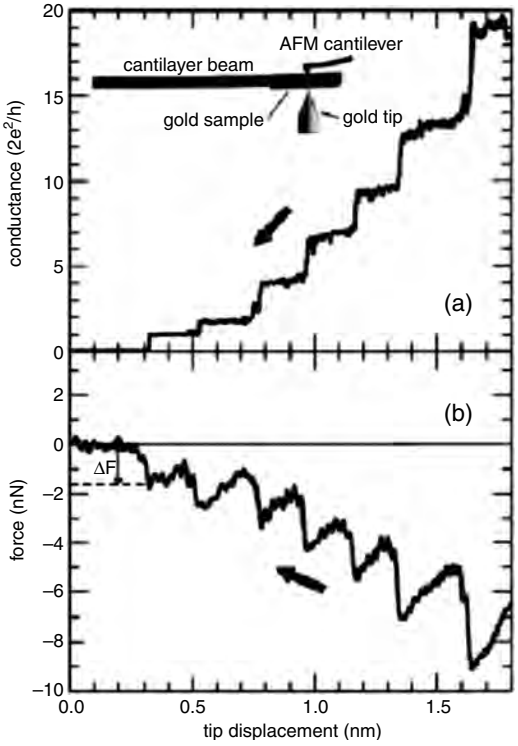


Figure 36. (a) Conductance and (b) force during the elongation of a gold nanowire. Reprinted with permission from [171], G. Rubio et al., *Phys. Rev. Lett.* 76, 2302 (1996). © 1996, American Physical Society.

where T_k is the transmission probability for the kth electron conduction channel. The number of channels is proportional to the contact area, and T_k is close to 1, so that the conductance G is proportional to the contact area.

Rubio et al. [171] brought into contact a tip and a sample of gold to form a single asperity contact. A small bias voltage was applied and the current through the contact was measured. Figure 36a shows the conductance of a tip, which was pulled back from the sample after the formation of the contact, and Figure 36b shows the corresponding force. The conductance changes stepwise, and each step is associated with a relaxation in the normal force F_N, which, otherwise, increases elastically. Each jump can be associated with an atomic rearrangement of the contact, which changes the number of conduction channels.

8. CONCLUSIONS

We have emphasized the variety of friction and wear processes occurring in different physical systems, which have been investigated with the atomic force microscope. The unique capability to detect and control forces between a sharp nanometer-sized tip and a surface could reveal processes occurring on surfaces in motion down to the atomic scale, and a lot of information could be extracted also at a fundamental level (load and velocity dependence of friction).

Considering that it is practically impossible to report all the results currently obtained by friction force microscopy around the world, we mainly focused on some original experiments, which, in our opinion, led to a better understanding of friction and wear processes on the nanometer scale. The scenario can be extended by further readings from books and review articles on this subject [61, 116, 172–175].

Even if the experiments discussed constitute a significant breakthrough in our understanding of tribology, much work is still to be done. Subjects like temperature dependence of friction or friction changes in phase transitions require further investigation. The subject of friction and wear on the nanometer scale is extremely wide, and we believe that the observation of nature and the needs of technology will suggest new exciting experiments in this field.

GLOSSARY

Amonton's law The Amonton's law states that friction is proportional to the applied load. It is verified in macroscopic friction experiments.

Atomic stick-slip Mechanism in which the friction force between a sharp tip and a crystal lattice repeatedly builds up and then quickly slips at each atomic site.

Coulomb's law of friction The Coulomb's law of friction states that kinetic friction is independent of the sliding velocity. It is approximately fulfilled in macroscopic friction experiments.

Derjaguin-Muller-Toporov (DMT) model Model for contact and friction which includes the effect of asperity adhesion forces. The range of action of adhesive forces is assumed larger than the elastic deformation.

Frenkel-Kontorova-Tomlinson model Model for friction which considers a layer of atoms connected to a rigid upper plate via springs. The layer feels a spatially periodic potential. The plate is driven by a constant velocity.

Johnson-Kendall-Roberts (JKR) model Model for contact and friction which includes the effect of asperity adhesion forces. The range of action of adhesive forces is assumed smaller than the elastic deformation.

Maugis-Dugdale model Model for contact and friction which includes the effect of asperity adhesion forces. Its range of applicability lies inbetween the JKR and the DMT model.

Quartz crystal microbalance Piezoelectric transducer which allows a mass change to be converted into a resonant frequency change.

Surface force apparatus Device used to measure the interaction force between two macroscopic surfaces as a function of surface separation.

Tomlinson model Model for friction which considers a single atom connected to a rigid upper plate via a spring. The atom feels a spatially periodic potential. The plate is driven by a constant velocity.

REFERENCES

1. G. Binnig, C. F. Quate, and Ch. Gerber, *Phys. Rev. Lett.* 56, 930 (1986).
2. R. Lüthi, E. Meyer, H. Haefke, L. Howald, W. Gutmannsbauer, M. Guggisberg, M. Bammerlin, and H.-J. Güntherodt, *Surf. Sci.* 338, 247 (1995).

3. H. Hertz, *J. Reine Angew. Math.* 92, 156 (1881).
4. Dowson, "History of Tribology," Longman, London, 1979.
5. J. A. Greenwood and J. B. P. Williamson, *Proc. Roy. Soc. London Ser. A* 295, 300 (1966).
6. B. N. J. Persson, *Phys. Rev. Lett.* 87, 116101 (2001).
7. E. Meyer, R. M. Overney, R. Lüthi, D. Brodbeck, L. Howald, J. Frommer, H.-J. Güntherodt, O. Wolter, M. Fujihira, H. Takano, and Y. Gotoh, *Thin Solid Films* 220, 132 (1992).
8. C. M. Mate, G. M. McClelland, R. Erlandsson, and S. Chiang, *Phys. Rev. Lett.* 59, 1942 (1987).
9. E. Gnecco, R. Bennewitz, T. Gyalog, Ch. Loppacher, M. Bammerlin, E. Meyer, and H.-J. Güntherodt, *Phys. Rev. Lett.* 84, 1172 (2000).
10. G. A. Tomlinson, *Philoso. Mag.* 7, 905 (1929).
11. E. Gnecco, R. Bennewitz, and E. Meyer, *Phys. Rev. Lett.* 88, 215501 (2002).
12. O. Marti, J. Colchero, and J. Mlynek, *Nanotechnology* 1, 141 (1990).
13. G. Meyer and N. Amer, *Appl. Phys. Lett.* 57, 2089 (1990).
14. G. Neubauer, S. R. Cohen, G. M. McClelland, D. E. Horn, and C. M. Mate, *Rev. Sci. Instrum.* 61, 2296 (1990).
15. G. M. McClelland and J. N. Glosli, *NATO ASI Ser. E* 220, 405 (1992).
16. R. Linnemann, T. Gotszalk, I. W. Rangelow, P. Dumania, and E. Oesterschulze, *J. Vac. Sci. Technol. B* 14, 856 (1996).
17. M. Nonnenmacher, J. Greschner, O. Wolter, and R. Kassing, *J. Vac. Sci. Technol. B* 9, 1358 (1991).
18. J. Cleveland, S. Manne, D. Bocek, and P. K. Hansma, *Rev. Sci. Instrum.* 64, 403 (1993).
19. J. L. Hutter and J. Bechhoefer, *Rev. Sci. Instrum.* 64, 1868 (1993).
20. H. J. Butt and M. Jaschke, *Nanotechnology* 6, 1 (1995).
21. J. M. Neumeister and W. A. Ducker, *Rev. Sci. Instrum.* 65, 2527 (1994).
22. D. F. Ogletree, R. W. Carpick, and M. Salmeron, *Rev. Sci. Instrum.* 67, 3298 (1996).
23. S. S. Sheiko, M. Möller, E. M. C. M. Reuvekamp, and H. W. Zandberger, *Phys. Rev. B* 48, 5675 (1993).
24. F. Atamny and A. Baiker, *Surf. Sci.* 323, L314 (1995).
25. J. S. Villarrubia, *J. Res. Natl. Inst. Stand. Technol.* 102, 425 (1997).
26. R. Staub, D. Alliata, and C. Nicolini, *Rev. Sci. Instrum.* 66, 2513 (1995).
27. S. M. Hues, C. F. Draper, K. P. Lee, and R. J. Colton, *Rev. Sci. Instrum.* 65, 1561 (1994).
28. J. Fu, *Rev. Sci. Instrum.* 66, 3785 (1995).
29. U. D. Schwarz, P. Köster, and R. Wiesendanger, *Rev. Sci. Instrum.* 67, 2560 (1996).
30. J. N. Israelachvili and D. Tabor, *Proc. R. Soc. London Ser. A* 331, 19 (1972).
31. J. N. Israelachvili, P. M. McGuiggan, and A. M. Homola, *Science* 240, 189 (1988).
32. B. Bhushan, J. N. Israelachvili, and U. Landman, *Nature* 374, 607 (1995).
33. P. Frantz, N. Agraït, and M. Salmeron, *Langmuir* 12, 3289 (1996).
34. S. J. O'Shea, M. E. Welland, and T. Rayment, *Appl. Phys. Lett.* 60, 2356 (1992).
35. S. J. O'Shea, M. E. Welland, and J. B. Pethica, *Chem. Phys. Lett.* 223, 336 (1994).
36. J. Krim and A. Widom, *Phys. Rev. B* 38, 12184 (1988).
37. J. Krim, D. H. Solina, and R. Chiarello, *Phys. Rev. Lett.* 66, 181 (1991).
38. F. P. Bowden and D. Tabor, "The Friction and Lubrication of Solids." Oxford Univ. Press, London, 1950.
39. L. D. Landau and E. M. Lifshitz, in "Introduction into Theoretical Physics," Vol. 7. Nauka, Moscow, 1998.
40. K. L. Johnson, K. Kendall, and A. D. Roberts, *Proc. Roy. Soc. London Ser. A* 324, 301 (1971).
41. B. V. Derjaguin, V. M. Muller, and Y. P. Toporov, *J. Colloid Interface Sci.* 53, 314 (1975).
42. D. Tabor, *J. Colloid Interface Sci.* 58, 2 (1977).
43. D. Maugis, *J. Colloid Interface Sci.* 150, 243 (1992).
44. U. D. Schwarz, O. Zwörner, P. Köster, and R. Wiesendanger, *Phys. Rev. B* 56, 6987 (1997).
45. E. Meyer, R. Lüthi, L. Howald, M. Bammerlin, M. Guggisberg, and H.-J. Güntherodt, *J. Vac. Sci. Technol. B* 14, 1285 (1996).
46. R. W. Carpick, N. Agraït, D. F. Ogletree, and M. Salmeron, *J. Vac. Sci. Technol. B* 14, 1289 (1996).
47. C. Polaczyk, T. Schneider, J. Schöfer, and E. Santner, *Surf. Sci.* 402, 454 (1998).
48. S. P. Jarvis, A. Oral, T. P. Weihs, and J. B. Pethica, *Rev. Sci. Instrum.* 64, 3515 (1993).
49. R. W. Carpick, D. F. Ogletree, and M. Salmeron, *Appl. Phys. Lett.* 70, 1548 (1997).
50. M. A. Lantz, S. J. O'Shea, A. C. Hoole, and M. E. Welland, *Appl. Phys. Lett.* 70, 970 (1997).
51. M. A. Lantz, S. J. O'Shea, M. E. Welland, and K. L. Johnson, *Phys. Rev. B* 55, 10776 (1997).
52. K. L. Johnson, "Contact Mechanics," Cambridge Univ. Press, Cambridge, UK, 1985.
53. M. Enachescu, R. J. A. van der Oetelaar, R. W. Carpick, D. F. Ogletree, C. F. J. Flipse, and M. Salmeron, *Phys. Rev. Lett.* 81, 1877 (1998).
54. M. Enachescu M, R. J. A. van den Oetelaar, R. W. Carpick, D. F. Ogletree, C. F. J. Flipse, and M. Salmeron, *Trib. Lett.* 7, 73 (1999).
55. Marti, B. Drake, and P. K. Hansma, *Appl. Phys. Lett.* 51, 484 (1987).
56. E. Meyer, R. M. Overney, L. Howald, R. Lüthi, J. Frommer, and H.-J. Güntherodt, *Phys. Rev. Lett.* 69, 1777 (1992).
57. B. J. Briscoe and D. C. B. Evans, *Proc. Roy. Soc. London Ser. A* 380, 389 (1982).
58. R. M. Overney, E. Meyer, J. Frommer, D. Brodbeck, R. Lüthi, L. Howald, H.-J. Güntherodt, M. Fujihira, H. Takano, and Y. Gotoh, *Nature* 359, 133 (1992).
59. C. D. Frisbie, L. F. Rozsnyai, A. Noy, M. S. Wrighton, and C. M. Lieber, *Science* 265, 2071 (1994).
60. J. B. D. Green, M. T. McDermott, and M. D. Porter, *J. Phys. Chem.* 99, 10960 (1995).
61. R. W. Carpick and M. Salmeron, *Chem. Rev.* 97, 1163 (1997).
62. M. Mate, *Wear* 168, 17 (1993).
63. R. Buzio, E. Gnecco, C. Boragno, and U. Valbusa, *Carbon* 40, 883 (2002).
64. S. S. Perry, C. M. Mate, R. L. White, and G. A. Somorjai, *IEEE Trans. Magn.* 32, 115 (1996).
65. R. Lüthi, E. Meyer, H. Haefke, L. Howald, W. Gutmannsbauer, and H.-J. Güntherodt, *Science* 266, 1979 (1994).
66. L. Scandella, E. Meyer, L. Howald, R. Lüthi, M. Guggisberg, J. Gobrecht, and H.-J. Güntherodt, *J. Vac. Sci. Technol. B* 14, 1255 (1996).
67. T. Teuschler, K. Mahr, S. Miyazaki, M. Hundhausen, and L. Ley, *Appl. Phys. Lett.* 66, 2499 (1995).
68. J. Tamayo, L. González, Y. González, and R. García, *Appl. Phys. Lett.* 68, 2997 (1996).
69. J. Tamayo, R. García, T. Utzmeier, and F. Briones, *Phys. Rev. B* 55, R13436 (1997).
70. R. Lüthi, H. Haefke, K.-P. Meyer, E. Meyer, and L. Howald, *J. Appl. Phys.* 74, 7461 (1993).
71. L. M. Eng, M. Friedrich, J. Fousek, and P. Günter, *J. Vac. Sci. Technol. B* 14, 1191 (1996).
72. H. Bluhm, U. D. Schwarz, and R. Wiesendanger, *Phys. Rev. B* 57, 161 (1998).
73. H. Bluhm, T. Inoue, and M. Salmeron, *Phys. Rev. B* 61, 7760 (2000).
74. A. L. Weisenhorn, P. Maivald, H.-J. Butt, and P. K. Hansma, *Phys. Rev. B* 45, 11226 (1992).

75. M. Binggeli and C. M. Mate, *Appl. Phys. Lett.* 65, 415 (1994).
76. A. Schumacher, N. Kruse, R. Prins, E. Meyer, R. Lüthi, L. Howald, H.-J. Güntherodt, and L. Scandella, *J. Vac. Sci. Technol. B* 14, 1264 (1996).
77. C. A. J. Putman, M. Igarshi, and R. Kaneko, *Appl. Phys. Lett.* 66, 3221 (1995).
78. A. Marti, G. Hähner, and N. D. Spencer, *Langmuir* 11, 4632 (1995).
79. G. Hähner, A. Marti, and N. D. Spencer, *Trib. Lett.* 3, 359 (1997).
80. V. N. Koinkar and B. Bhushan, *J. Appl. Phys.* 81, 2472 (1999).
81. S. Sundararajan and B. Bhushan, *J. Appl. Phys.* 88, 4825 (2000).
82. D. Tomanek, W. Zhong, and H. Thomas, *Europhys. Lett.* 15, 887 (1991).
83. J. Colchero, A. M. Baro, and O. Marti, *Tribol. Lett.* 2, 327 (1996).
84. W. Zhong and D. Tomànek, *Phys. Rev. Lett.* 64, 3054 (1990).
85. T. Gyalog, M. Bammerlin, R. Lüthi, E. Meyer, and H. Thomas, *Europhys. Lett.* 31, 269 (1995).
86. N. Sasaki, M. Kobayashi, and M. Tsukada, *Surf. Sci.* 358, 92 (1996).
87. N. Sasaki, M. Kobayashi, and M. Tsukada, *Phys. Rev. B* 54, 2138 (1996).
88. H. Hölscher, U. D. Schwarz, and R. Wiesendanger, *Europhys. Lett.* 36, 19 (1996).
89. H. Hölscher, U. D. Schwarz, and R. Wiesendanger, *Surf. Sci.* 375, 395 (1997).
90. H. Hölscher, U. D. Schwarz, O. Zwörner, and R. Wiesendanger, *Phys. Rev. B* 57, 2477 (1998).
91. K. L. Johnson, J. Woodhouse, *Trib. Lett.* 5, 155 (1998).
92. E. Gnecco, R. Bennewitz, and E. Meyer, *J. Phys.: Condens. Matter* 13, R619 (2001).
93. T. Gyalog, M. Bammerlin, R. Lüthi, E. Meyer, and H. Thomas, *Europhys. Lett.* 31, 269 (1995).
94. M. Weiss and F. J. Elmer, *Phys. Rev. B* 53, 7539 (1996).
95. T. Gyalog and H. Thomas, *Europhys. Lett.* 37, 195 (1997).
96. U. Landman, W. D. Luedtke, and M. W. Ribarsky, *J. Vac. Sci. Technol. A* 7, 2829 (1989).
97. U. Landman, W. D. Luedtke, and E. M. Ringer, *Wear* 153, 3 (1992).
98. J. A. Harrison, C. T. White, R. J. Colton, and W. Brenner, *Surf. Sci.* 271, 57 (1992).
99. M. R. Sørensen, K. W. Jacobsen, and P. Stoltze, *Phys. Rev. B* 53, 2101 (1996).
100. A. L. Shluger, A. L. Rohl, R. M. Wilson, and R. T. Williams, *J. Vac. Sci. Technol. B* 13, 1155 (1995).
101. A. L. Shluger, A. L. Rohl, R. T. Williams, and R. M. Wilson, *Phys. Rev. B* 52, 11398 (1995).
102. A. L. Shluger, R. T. Williams, and A. L. Rohl, *Surf. Sci.* 343, 273 (1995).
103. A. I. Livshitz and A. L. Shluger, *Phys. Rev. B* 56, 12482 (1997).
104. T. Ohzono and M. Fujihira, *Phys Rev. B* 62, 17055 (2000).
105. T. Ohzono and M. Fujihira, *Trib. Lett.* 9, 63 (2000).
106. A. Buldum and S. Ciraci, *Phys. Rev. B* 57, 2468 (1998).
107. R. Komanduri, N. Chandrasekaran, and L. M. Raff, *Phys. Rev. B* 61, 14007 (2000).
108. T. H. Fang, C. I. Weng, and J. G. Chang, *Surf. Sci.* 501, 138 (2002).
109. S. Ciraci, E. Tekman, A. Baratoff, and I. Batra, *Phys. Rev. B* 46, 10411 (1992).
110. M. Sørensen, K. Jacobsen, and H. Jonsson, *Phys. Rev. Lett.* 77, 5067 (1996).
111. R. Perez, M. Payne, and A. Simpson, *Phys. Rev. Lett.* 75, 4748 (1995).
112. R. Astala, M. Kaukinen, R. M. Nieminen, and T. Heine, *Phys. Rev. B* 61, 2973 (2000).
113. J. B. Pethica, *Phys. Rev. Lett.* 57, 3235 (1986).
114. L. Howald, E. Meyer, R. Lüthi, H. Haefke, R. Overney, H. Rudin, and H.-J. Güntherodt, *Appl. Phys. Lett.* 63, 117 (1993).
115. R. Lüthi, E. Meyer, M. Bammerlin, L. Howald, H. Haefke, T. Lehmann, C. Loppacher, H.-J. Güntherodt, T. Gyalog, and H. Thomas, *J. Vac. Sci. Technol. B* 14, 1280 (1996).
116. E. Meyer, R. M. Overney, K. Dransfeld, and T. Gyalog, "Nanoscience: Friction and Rheology on the Nanometer Scale." World Scientific, Singapore, 1998.
117. L. Howald, R. Lüthi, E. Meyer, and H.-J. Güntherodt, *Phys. Rev. B* 51, 5484 (1995).
118. R. Bennewitz, T. Gyalog, M. Guggisberg, M. Bammerlin, E. Meyer, and H.-J. Güntherodt, *Phys. Rev. B* 60, R11301 (1999).
119. R. Bennewitz, E. Gnecco, T. Gyalog, and E. Meyer, *Trib. Lett.* 10, 51 (2001).
120. G. J. Germann, S. R. Cohen, G. Neubauer, G. M. McClelland, and H. Seki, *J. Appl. Phys.* 73, 163 (1993).
121. R. J. A. van den Oetelaar and C. F. J. Flipse, *Surf. Sci.* 384, L828 (1997).
122. S. Morita, S. Fujisawa, and Y. Sugawara, *Surf. Sci. Rep.* 23, 3 (1996).
123. S. Fujisawa, E. Kishi, Y. Sugawara, and S. Morita, *Phys. Rev. B* 51, 7849 (1995).
124. S. Fujisawa, E. Kishi, Y. Sugawara, and S. Morita, *Nanotechnology* 5, 8 (1994).
125. H. Kawakatsu and T. Saito, *J. Vac. Sci. Technol. B* 14, 872 (1996).
126. M. Hirano, K. Shinjo, R. Kaneko, and Y. Murata, *Phys. Rev. Lett.* 67, 2642 (1991).
127. M. Hirano, K. Shinjo, R. Kaneko, and Y. Murata, *Phys. Rev. Lett.* 78, 1448 (1997).
128. R. M. Overney, H. Takano, M. Fujihira, W. Paulus, and H. Ringsdorf, *Phys. Rev. Lett.* 72, 3546 (1994).
129. M. Liley, D. Gourdon, D. Stamou, U. Meseth, T. M. Fischer, C. Lautz, H. Stahlberg, H. Vogel, N. A. Burnham, and C. Duschl, *Science* 280, 273 (1998).
130. P. E. Sheehan and C. M. Lieber, *Science* 272, 1158 (1996).
131. M. R. Falvo, J. Steele, R. M. Taylor II, and R. Superfine, *Trib. Lett.* 9, 73 (2000).
132. O. Zwörner, H. Hölscher, U. D. Schwarz, and R. Wiesendanger, *Appl. Phys. A* 66, S263 (1998).
133. D. Gourdon, N. A. Burnham, A. Kulik, E. Dupas, F. Oulevey, G. Gremaud, D. Stamou, M. Liley, Z. Dienes, H. Vogel, and C. Duschl, *Trib. Lett.* 3, 317 (1997).
134. T. Bouhacina, J. P. Aimé, S. Gauthier, D. Michel, and V. Heroguez, *Phys. Rev. B* 56, 7694 (1997).
135. H. J. Eyring, *J. Chem. Phys.* 3, 107 (1937).
136. J. N. Glosli and G. M. McClelland, *Phys. Rev. Lett.* 70, 1960 (1993).
137. E. Riedo, F. Lévy, and H. Brune, *Phys. Rev. Lett.* 88, 185505 (2002).
138. M. He, A. S. Blum, G. Overney, and R. M. Overney, *Phys. Rev. Lett.* 88, 154302 (2002).
139. Ch. Loppacher, R. Bennewitz, O. Pfeiffer, M. Guggisberg, M. Bammerlin, S. Schär, V. Barwich, A. Baratoff, and E. Meyer, *Phys. Rev. B* 62, 13674 (2000).
140. J. P. Aimé, R. Boisgard, L. Nony, and G. Couturier, *Phys. Rev. Lett.* 82, 3388 (1999).
141. M. Gauthier and M. Tsukada, *Phys. Rev. B* 60, 11716 (1999).
142. W. Denk and D. W. Pohl, *Appl. Phys. Lett.* 59, 2171 (1991).
143. S. Hirsekorn, U. Rabe, A. Boub, and W. Arnold, *Surf. Interface Anal.* 27, 474 (1999).
144. U. Dürig, *NATO ASI Ser. E* 286, 353 (1995).
145. N. Sasaki and M. Tsukada, *Jpn. J. Appl. Phys.* 39, L1334 (2000).
146. B. Gotsmann and H. Fuchs, *Phys. Rev. Lett.* 86, 2597 (2001).
147. O. Pfeiffer, R. Bennewitz, A. Baratoff, E. Meyer, and P. Grütter, *Phys. Rev. B* 65, 161403 (2002).
148. H. J. Hug, and A. Baratoff, in "Noncontact Atomic Force Microscopy," (S. Morita, R. Wiesendanger, and E. Meyer, Eds.), p. 395. Springer, Berlin, 2002.

149. B. N. J. Persson, *Phys. Rev. B* 44, 3277 (1991).
150. C. Mak, C. Daly, and J. Krim, *Thin Solid Films* 253, 190 (1994).
151. A. Dayo, W. Alnasrallah, and J. Krim, *Phys. Rev. Lett.* 80, 1690 (1998).
152. B. L. Mason, *Trib. Lett.* 10, 59 (2001).
153. J. Hu, X. D. Xiao, D. F. Ogletree, and M. Salmeron, *Surf. Sci.* 327, 358 (1995).
154. B. Bhushan and V. N. Koinkar, *Appl. Phys. Lett.* 64, 1653 (1994).
155. B. Bhushan and A. V. Kulkarni, *Thin Solid Films* 278, 49 (1996).
156. J. Zhaoguo, C. J. Lu, D. B. Bogy, C. S. Bhatla, and T. Myamoto, *Thin Solid Films* 258, 75 (1995).
157. S. Miyake, *Appl. Phys. Lett.* 65, 980 (1994).
158. S. Miyake, *Appl. Phys. Lett.* 67, 2925 (1995).
159. A. Khurshudov and K. Kato, *Ultramicroscopy* 60, 11 (1995).
160. A. Khurshudov and K. Kato, *Trib. Lett.* 2, 345 (1996).
161. K. Herrmann, N. M. Jennett, W. Wegener, J. Meneve, K. Hasche, R. Seemann, *Thin Solid Films* 377, 394 (2000).
162. B. Bhushan, A. V. Kulkarni, W. Bonin, and J. T. Wyrobek, *Philos. Mag.* 74, 1117 (1996).
163. B. Bhushan, *Wear* 251, 1105 (2001).
164. B. Irmer, R. H. Blick, F. Simmel, W. Gödel, H. Lorenz, and J. P. Kotthaus, *Appl. Phys. Lett.* 73, 2051 (1998).
165. J. Cortes Rosa, M. Wendel, H. Lorenz, J. P. Kotthaus, M. Thomas, and H. Kroemer, *Appl. Phys. Lett.* 73, 2684 (1998).
166. S. Kopta and M. Salmeron, *J. Chem. Phys.* 113, 8249 (2000).
167. J. T. Dickinson, N. S. Park, M. W. Kim, and S. C. Langford, *Trib. Lett.* 3, 69 (1997).
168. S. Nakahara, S. C. Langford, and J. T. Dickinson, *Trib. Lett.* 1, 277 (1995).
169. T. Thundat, B. C. Sales, B. C. Chakoumakos, L. A. Boatner, D. P. Allison, and R. J. Warmack, *Surf. Sci.* 293, L863 (1993).
170. M. Buttiker, Y. Imry, R. Landauer, and S. Pinhas, *Phys. Rev. B* 31, 6207 (1985).
171. G. Rubio, N. Agrait, and S. Vieira, *Phys. Rev. Lett.* 76, 2302 (1996).
172. "Handbook of Micro/Nanotribology" (B. Bhushan, Ed.). CRC Press, Boca Raton, FL, 1999.
173. B. N. J. Persson and E. Tosatti, "Physics of Sliding Friction." Kluwer, Dordrecht, 1996.
174. B. N. J. Persson, "Sliding Friction," Springer-Verlag, Berlin, 1998.
175. G. V. Dedkov, *Phys. Status Solid* 179, 3 (2000).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanotribology of Carbon Films

F. L. Freire Jr., R. Prioli

Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, Brazil

CONTENTS

1. INTRODUCTION

Amorphous carbon films are made of clusters of sp^2-hybridized carbon atoms, with typical size of 1 nm, connected to each other by sp^3-hydridized carbon atoms. The electronic and optical properties of these materials are mainly controlled by the size of the sp^2 clusters, while the mechanical properties are given by the degree of interconnectivity of the amorphous skeleton, that is, the fraction of sp^3-hybridized carbon atoms present in the matrix [1]. Film properties can be tuned by choosing the deposition technique with the appropriate deposition conditions. Among the different deposition parameters, the energy of the impinging ions plays the most important role and controls the sp^2/sp^3-hybridized carbon atoms fraction. In this way, films with properties typical of polymers, *polymer-like films*, or of graphite or diamond, *graphite-like* and *diamond-like films*, respectively, can be synthesized. On the other hand, the incorporation of elements such as nitrogen, silicon, or fluorine can optimize some of these properties [2] or even induce the formation of new nanostructured materials, such as the *fullerene-like* carbon films, films with strongly interacting curved graphene planes. The fullerene-like structure results from curving of the graphene planes induced by nitrogen incorporation and is similar to that found in carbon nanotubes and fullerenes [3].

The study of the nanometer scale of the physical phenomena related to the interaction of surfaces in contact and in relative motion, nanotribology, was made possible with the invention of the atomic force microscope in 1986 [4]. These phenomena are extremely important in situations like manipulation of atoms and molecules in surfaces, as well as in the operation of electromechanical devices. For example, it was recently demonstrated that the knowledge of the energy dissipation mechanism when an atomic force microscope tip scratches the surface of a hard film might be useful for surface chemical and mechanical modifications with high spatial resolution [5]. For complete understanding of the phenomena involved, basic studies on the friction laws in nanoscale are necessary. In the last decade, with the development of friction force microscopy, the study of wear and friction properties of the materials has received increasing attention [6, 7]. Despite this, the study of the lubricant and tribological properties of carbon-based films and other nanostructured materials is still in the beginning stages and deserves more work for a complete description of the tribological processes on the nanometer scale.

In the next sections, we will briefly review present knowledge on the properties and structure of amorphous carbon films as well as current applications. We will follow with a short discussion of the basic concepts involved in both atomic force microscopy and friction force microscopy. After that, we will discuss recent results on the tribological properties of carbon-based films. We will devote special attention to the effects of the incorporation of hydrogen and nitrogen into amorphous carbon (a-C) films.

2. AMORPHOUS CARBON FILMS

Carbon films are presently being used in a wide variety of applications. In particular, films deposited by magnetron sputtering, plasma-enhanced chemical vapor deposition (PECVD) and filtered cathodic vacuum arc (FCVA) are used as protective overcoats for automotive components, shaving blades, biomedical implants, and computer hard disks [1]. The use of hydrogenated carbon films as the interconnect dielectric in ultra-large scale integrated chips was also proposed [8]. In addition to these many applications, a-C films are materials with considerable interest from an intrinsically basic point of view, and they have been the subject of intensive research in the last three decades.

Since the pioneering work of Aisenberg and Chabot [9], hyperthermal (energies ~1–1000 eV) carbon-containing

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (899–917)

species have been extensively used for the synthesis of a-C films [10]. The subplantation model suggests that the deposition of hyperthermal species is a shallow implantation process, where the incorporation of carbon species in the subsurface layers followed by large internal stresses is the dominant mechanism, responsible for the formation of a sp^3 carbon rich dense phase, the diamond-like phase [11–13]. The energy of the impinging species plays the main role in the control of the carbon bonding hybridization. In fact, the evolution of the dense phase in a-C films occurs by a preferential displacement mechanism. It occurs when the incident ion energy is sufficient for the displacement of atoms with low displacement energy, E_d, the sp^2-hybridized carbon atoms, but it is insufficient for the displacement of high E_d atoms, the sp^3-hybridized carbon atoms. In subplantation, the increased density stabilizes sp^3 hybridization over sp^2 hybridization. For some experimental conditions, a-C films with up to 80% of carbon atoms with sp^3 hybridization can be obtained. These films are called tetrahedral amorphous carbon (ta-C) films and can be hydrogenated or not. For films deposited by FCVA using C^+ beams, the maximum fraction of sp^3 hybridization occurs at around 100 eV [14], which is in good agreement with subplantation model predictions [11].

Amorphous carbon films present important properties that are closely related to the film microstructure [15]. Following Robertson [1], the sp^3 bonds are responsible for the high hardness, which can reach ~80 GPa (Young's modulus of 700–800 GPa) in ta-C films grown by filtered arc and laser ablation methods, and because sp^2 bonds are weaker than sp^3 bonds, the respective π states of these bonds lie closer to the Fermi level, controlling the gap size and the optical properties of the materials. By choosing the deposition technique and the appropriate deposition parameters, the ratio between carbon atoms in sp^2 and sp^3 hybridization can be tuned and the film properties can be controlled.

The incorporation of hydrogen in the amorphous skeleton also plays an important role, as is illustrated in Figure 1. In this figure, a ternary phase diagram of the amorphous carbon–hydrogen system is presented. Besides the 100% sp^3 (diamond) and 100% sp^2 (graphite) phases, a-C films with the fraction of sp^3 bonding in the range between 5 and 80% can be produced. The hydrogen content of these films can vary from 0 to 50 at.%. The amorphous hydrogenated carbon (a-C:H) films can be deposited by both sputtering and PECVD techniques in a large range of sp^2/sp^3-hybridized carbon atom ratios. Depending on the deposition conditions, substrate-bias voltage, substrate temperature, pressure, and gas precursor atmosphere, they can be either diamond-like (sp^3-rich films) or graphitic-like (sp^2-rich films). However, a-C:H films with high H content (~50 at.%) and a high fraction of carbon atoms with sp^3 hybridization have a polymeric character. The region of larger H content is forbidden, because this corresponds to molecules, not a continuous network solid. Recently, it was shown that the subplantation model could be applied to describe the mechanisms of a-C:H growth by PECVD. In fact, the internal stress and hardness of a-C:H films deposited by radio frequency (rf)-PECVD in pure methane atmosphere show a dependence with the self-bias that is well described by the subplantation model. In this case, hardness and stress presenting

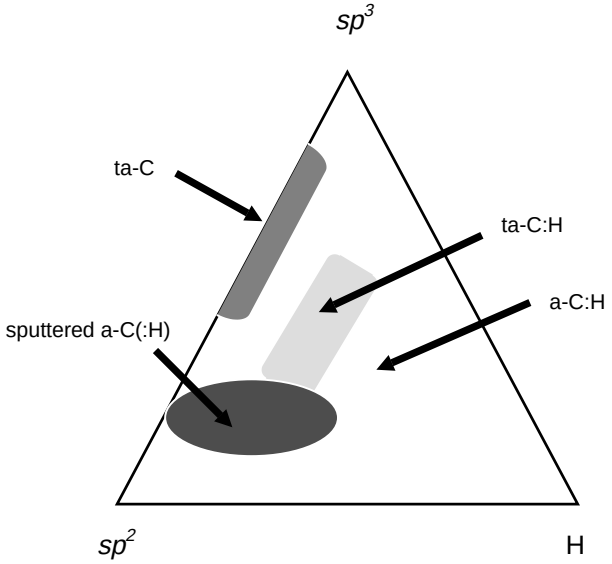


Figure 1. Ternary phase diagram of the amorphous carbon–hydrogen system. Reprinted with permission from [63], J. R. Robertson, *Thin Solid Films* 383, 81 (2001). © 2001, Elsevier Science.

maxima at around −100 V of self-bias voltage, V_b, for films deposited at 1 Pa [16]. At this pressure the energy of the impinging ions is essentially eV_b, where e is the electron charge.

One way of changing a-C:H film properties is through the incorporation of different elements such as N, F, and Si during the film growth. The incorporation of Si, for example, stabilizes sp^3 carbon bonds and improves the thermal stability of a-C:H films [17]. The incorporation of fluorine in a-C:H films has also attracted interest in the last few years. Despite the remarkable lubricant properties of poly(tetrafluoroethylene), recent research on fluorine incorporation into a-C:H films is mainly motivated by its electrical characteristics. In fact, to improve the switching performance of future ultra-large scale integrated circuits, insulator films with dielectric constants lower than that of SiO_2 are needed to reduce the capacitance of interlayer insulators [8]. Fluorinated amorphous carbon (a-C:H:F) films have been proposed as possible candidates due to their low dielectric constant [18–22]. These films can be plasma deposited in a wide compositional range, the film fluorine content being determined primarily by the precursor atmosphere [19]. It has been found that the dielectric constant decreases with an increase in the F/H ratio of the precursor atmosphere for films deposited by PECVD [19]. However, the poor thermal stability of a-C:H:F films inhibits their application as a low K dielectric in semiconductor devices, because important thickness reduction was observed for annealing temperatures higher than 300°C [22]. Notwithstanding the interest in the dielectric properties of fluorinated amorphous carbon films, not much attention has been paid to the effect of fluorine incorporation on the mechanical and tribological properties of a-C:H films [23–26]. Similarly, the effect of ionic bombardment during a-C:H:F film growth has not been investigated in great detail [25, 27]. The incorporation of fluorine increases the film surface energy, and because

one makes the choice of the appropriate set of deposition parameters a hard, anti-adhesive, and low friction coating can be deposited with a high deposition rate [28].

In the last few years, an important portion of the research effort on a-C:H films was dedicated to the study of the nitrogen incorporation [29, 30]. The main purpose of this research effort was the intention to synthesize the β-C_3N_4 solid, proposed by Liu and Cohen [31], structurally analogous to β-silicon nitride (sp^3-carbon atoms bonded to sp^2-nitrogen) and with mechanical properties comparable to those of crystalline diamond. Many attempts to synthesize β-C_3N_4 samples have failed: high pressure–high temperature techniques produced only graphitic materials [32, 33], and the film deposition techniques usually produce amorphous films [34, 35]. In fact, only in a few cases was the presence of some small crystalline grains embedded in an amorphous matrix reported [36, 37]. In these cases, an electron diffraction pattern somehow related to the predicted one was obtained from those grains. The difficulties in synthesizing β-C_3N_4 were tentatively explained by molecular dynamics calculations showing that strong nitrogen incorporation into the amorphous carbon network catalyzes carbon undercoordination, which in turn causes the nitrogen to develop CN double and triple bonds [38]. Electron energy loss spectroscopy measurements support these calculations, because they show that nitrogen incorporation into amorphous carbon films, as also observed in amorphous hydrogenated carbon films [39], induces an increase in the fraction of carbon atoms presenting sp^2-hybridization [40]. Results provided by neutron and X-ray diffraction experiments also revealed that nitrogen incorporation induces a transformation of carbon hybridization from sp^3 to sp^2 states, with the presence of both C═N and C≡N bonds [41, 42].

Nitrogen incorporation into a-C:H films was found to modify the structure and properties of these films. Concerning the mechanical properties, a reduction on the internal compressive stress was observed, with minor changes in the mechanical hardness [43–45]. Concerning the modification of electrical and optical properties, it was found that nitrogen could electronically dope a-C:H films with the simultaneous reduction of the electronic defect density. This makes possible the use of a-C:H:N films as a semiconductor material [46, 47].

In an attempt to avoid the metastability of β-C_3N_4, a new approach was tried: the growth of superlattices involving very thin layers, a few nanometers thick, of carbon nitride and transition metal nitrides, the last one providing a lattice-matched structural template to seed the growth of the β-C_3N_4 phase. The production of a hard (~50 GPa) and fully crystalline carbon nitride composite material was claimed to be achieved [48, 49]. In a recent publication, a detailed structural characterization of this material was performed, but, unfortunately, despite the fact that the presence of sp^3-hybridized carbon was revealed by several analytical techniques, Raman and electron energy loss spectroscopies [50], it was not possible to clearly determine the presence of the β-C_3N_4 phase.

Despite the difficulties in synthesizing crystalline β-C_3N_4 or any other crystalline structure of carbon nitride, amorphous carbon–nitrogen films, a-CN_x, especially those deposited by reactive sputtering, are very interesting because of their application as protective coatings. In fact, a-CN_x films with hardness of ~20 GPa have been obtained by using different configurations of sputtering deposition systems: direct current (dc)- or rf-magnetron, ionized magnetron, and unbalanced dc-magnetron [51–53]. The substrate bias is the relevant parameter to optimize the mechanical and tribological properties of the films: a negative substrate bias is necessary for the production of hard a-CN_x films with improved tribological properties. The results of hard disk durability tests show that carbon–nitrogen coatings have greater wear resistance than pure argon-sputtered carbon films [54, 55]. As a consequence, a-CN_x films are currently being used as the protective coating on many commercial hard magnetic disk systems [56]. The effects of the substrate temperature on the film microstructure were also investigated. As a general rule, films deposited at low temperatures onto a floating substrate are amorphous, and the nitrogen content and the deposition rate decrease with increasing deposition temperature. Cuomo and collaborators [57], based on infrared (IR) absorption spectroscopy results, proposed a structure for a-CN_x films deposited by rf-sputtering onto a floating substrate that they called paracyanogen-like, a solid polymer form of cyanogen, —$(CN)_n$—. This material is formed by aromatic rings with C—N and C═N bonds, connected to hydrogen or CN radicals.

On the other hand, high-resolution transmission electron microscopy experiments performed on a-CN_x films deposited at temperatures higher than 250 °C revealed the presence of distorted graphite-like structures consisting of buckled and curved basal planes, giving rise to a fullerene-like microstructure [58]. Fullerene-like carbon (FLC) films with strongly interacting graphene planes have the unique properties of being hard and extremely elastic [3, 58]. The curving of graphene planes facilitates the formation of sp^3 bonds between intersecting graphene planes, resulting in a predominantly covalent sp^2 network. The highly elastic nature (up to 90% elastic recovery after indentation) of the FLC films makes it difficult to extract hardness values with the nanoindentation technique. However, the Young's modulus of FLC thin films measured using the surface acoustic wave method is approximately 500 GPa [3]. FLC films can be deposited using magnetron sputtering [58], laser arc [59], and localized high-pressure arc discharge (LHPA) [3]. In magnetron sputtering, high growth temperature (for formation of graphene sheets) and approximately 15% nitrogen in an $Ar + N_2$ atmosphere are required for the formation of the fullerene phase [58]. The presence of nitrogen in a hexagonal graphite lattice favors the formation of pentagons, inducing curvature in the graphene planes. In the laser arc and LHPA techniques, the *fullerene-like* structure results from the covalent interlinking of large fullerenes, nanotubes, and nanoparticles emitted from the cathode during the discharge.

Concerning the application of a-C films as tribological coatings, the protection of the hard magnetic disk deserves special attention, because it represents a unique combination of the latest technologies in tribology, thin film deposition, fluid mechanics, magnetism, and materials science. A schematic view of a hard magnetic disk is presented in Figure 2. With the recent development of magnetic

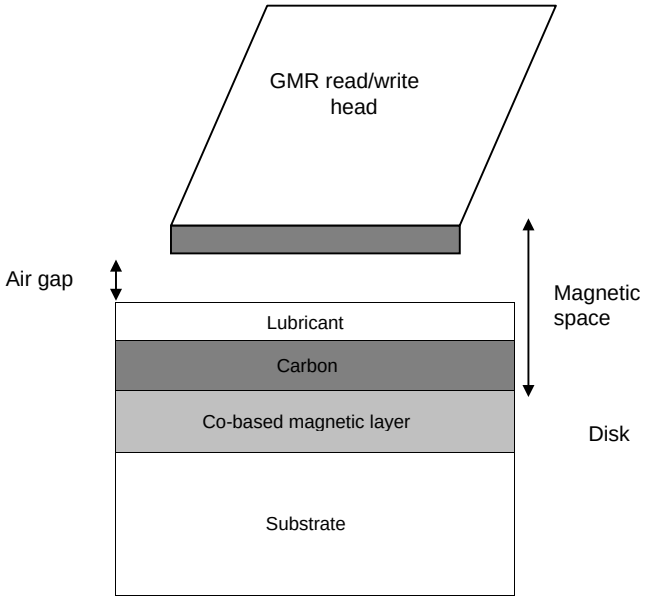


Figure 2. Schematic diagram of a hard magnetic disk and head.

devices, specially after the introduction of the magneto-resistive readers, the areal bit density in hard disks increased more than 60% a year, density that increased even faster in the last few years after the introduction of the giant magneto-resistive read technology. Today, even using a conservative approach, the goal of 100 Gbit/inch^2 could be achieved in less than 4 years from today [60]. As the recording density grows, reader stripe dimensions are vanishing. Head/magnetic media spacing must be reduced to achieve those densities. In fact, coatings that are around 1 nm thick will be needed, as shown in Figure 3, together with very restrictive conditions on surface roughness and lubricant layers. The spacing between the slider and the disk is steadily reduced and the disc revolutions per minute is heading toward 20,000 and beyond [61]. With the current load–unload technology, the overcoat is expected to provide some tribological protection together with protection against corrosion that requires a dense coating with low defect density. Under these conditions, drive performance improvement has to be necessarily combined with mechanical and tribological durability.

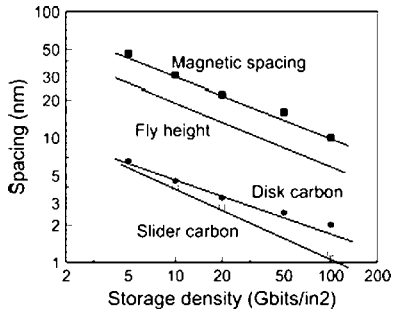


Figure 3. Variation of the carbon coating thickness on disk and sliders, the magnetic space, and the fly height with storage density. Reprinted with permission from [63], J. R. Robertson, *Thin Solid Films* 383, 81 (2001). © 2001, Elsevier Science.

Together with an important effort to extend to lower thickness the current coating technologies, that is, carbon nitride films deposited by pulsed magnetron sputtering [62], a lot of effort has been devoted to preparing a new generation of coatings, the ta-C films, based on mass selected ion deposition techniques [63], which are supposed to be able to fulfill all the requirements to achieve the goal of 100 Gbits/inch^2.

3. FORCE MICROSCOPY

3.1. Atomic Force Microscopy

A variety of instruments can be used to study friction, but it was after the invention of atomic force microscope (AFM)/friction force microscope (FFM) in the late 1980s that the field of tribology has gained new impulse. The force microscope has made possible the study of very low friction forces in a precise way. The instrument working principle of the AFM/FFM is very simple: a small tip is brought into contact with a surface, in a very controlled way, and scanned over the surface. While scanning, the normal and lateral movements of the microscope tip can be acquired. The normal movement of the tip is related by Hook's law to the normal force while its lateral movement, caused by the moment of the friction force acting on the tip–surface interface, is related to the friction forces.

In Figure 4, a typical AFM/lateral force microscope (LFM) scheme is presented. Usually, the samples are placed on the top of a piezoelectric tube ceramics that is going to control the lateral scan and the normal movements of the sample related to the tip. When scanning, the tip movement is measured by a position-sensitive photodetector. Although the top minus bottom (T – B) output from the detector is proportional to the tip normal displacement, the left minus right signal (L – R) is proportional to the torsion of the AFM tip. To keep the normal forces between the AFM tip and the sample constant during the scan, the T – B signal is compared with a reference value. Its difference, called error signal, is then used by the microscope feedback system to

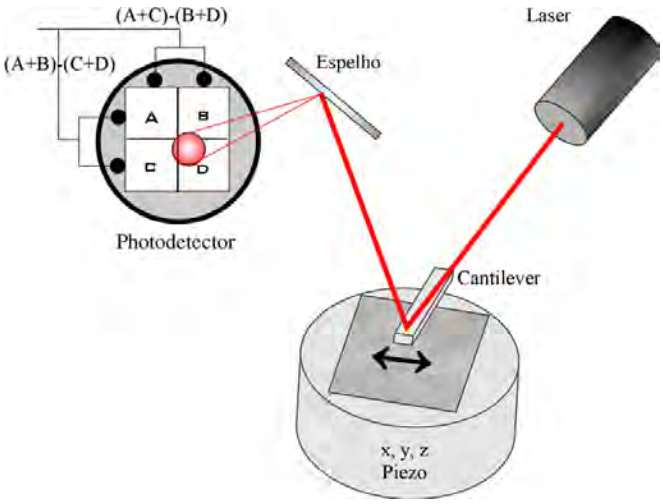


Figure 4. Schematic diagram of an atomic force microscope.

correct the distance between the tip and the sample, keeping the normal force constant during the whole image. If the feedback system is turned off, the piezoelectric ceramics are no longer moved in the normal direction, keeping the sample's height nearly constant. On this operational mode, the cantilever deflections are recorded as a function of the sample position. These two operational modes are then called constant force and constant height modes and have in common the fact that the total normal force, acting between the tip and surface atoms, is repulsive. In general, images with "low" resolution or large scan areas are performed in the force constant mode while high resolution images are acquired in the constant height mode. In Figure 5 we present two examples of these extreme cases and show an image obtained from an a-C film deposited by sputtering and an image obtained from a highly oriented pyrolytic graphite (HOPG) surface.

The AFM also has two additional modes that are called the noncontact and tapping mode. In both cases, instead of acquiring the direct bending of the cantilever tip, the amplitude, phase, or frequency from a mechanically vibrated tip is acquired. This signal is again compared with a reference value and its error is used by the feedback system to correct the sample height. While in the tapping mode the force sensed by the vibrated tip ranges from attractive to repulsive, in the noncontact mode the forces are predominately attractive.

A typical force curve obtained in an AFM is presented in Figure 6. In such an experiment, the sample is moved upward or downward with the use of the piezoelectric ceramics while the cantilever deflection is acquired. In region 1, the tip is far from the sample surface, thus no effective force is sensed by the AFM tip. As the sample is moved upward, that is, the distance between tip and sample decreases, the cantilever tip bends toward the sample (region 2), indicating that the forces between them are attractive. As the sample keeps moving upward, the tip comes close to the surface so that the net interaction force becomes repulsive (region 3). As the sample is moved downward, the tip remains in contact with the sample surface, due to its adhesion to the surface (region 4), until the force at the cantilever overcomes the tip–surface adhesion. The reader has to note that the attractive forces between tip and surface during its approach are caused mainly by Van der Waals forces, the repulsive forces are caused by the repulsion of the tip–surface atom's electronic clouds, and finally, the tip–surface adhesion is caused mainly by the capillary condensation of water between the tip and the surface.

As mentioned above, the tip movement can be translated into forces by Hook's law. Therefore, the calibration of the AFM cantilever and its displacement sensor for friction forces is essential for the successfully measurement of friction forces. First, the elastic normal bending and torsion constants of the cantilever have to be calculated, which is usually done with measurement of the cantilever's

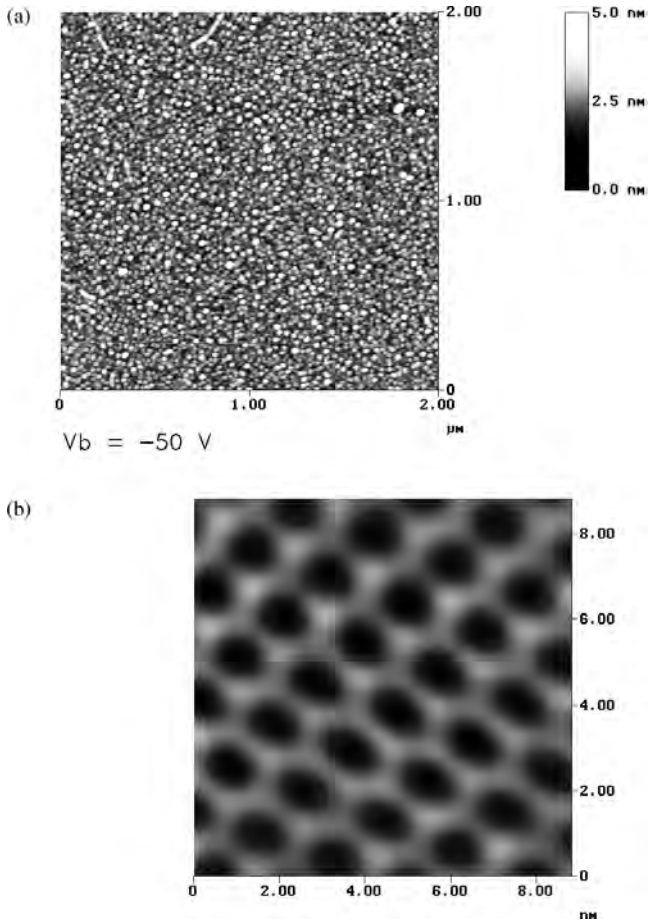


Figure 5. (a) Amorphous carbon film deposited by sputtering using bias voltage of −50 V and (b) boric acid single crystal (H_3BO_3) images obtained by the AFM.

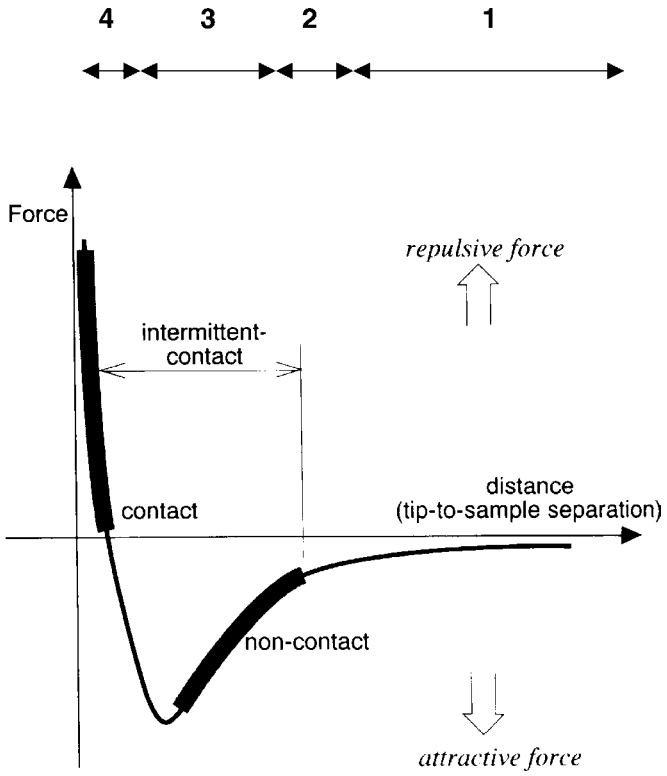


Figure 6. Typical force curve obtained with an AFM.

dimension by electronic microscopy and the use of a well-known static mechanics procedure [64]. Then, the displacement sensor can be calibrated to translate the normal and torsion bending angles in normal and friction forces [65].

3.2. Friction Force Microscopy

The energy that is dissipated at the interface when two materials are brought into contact and move with respect to each other is determined by the friction forces acting at the materials' interface. As defined, friction is always regarded as the force that opposes the movement of the surfaces. It was studied a long time ago by prominent scientists such as Leonardo da Vinci, Guillaume Amonton, and Charles A. Coulomb, among others [66]. They stated that friction is independent of the surface's contact area, that it is proportional to the normal forces acting between the surfaces, and that the kinetic friction is independent of the surface's relative velocity. Unfortunately, these laws seem to be no longer valid.

As observed by Mate et al. [67], the lateral movement of the scanning tip on a surface is shown to follow an atomic scale stick and slip behavior. The result obtained from a HOPG surface is shown in Figure 7. Here it was shown how the frictional force behaves during the stick and slip process as the surface is rastered in the xy plane. In this experience the sample is moved back and forth with the velocity of 40 nm/s in the x direction and 0.2 nm/s in the y direction. The critical friction force before the first slip occurs varies, depending on the position of the tip with respect to the graphite lattice. The slips occur where the image suddenly changes from bright to dark. Figure 7 clearly illustrates how the atomic structure of the surfaces influences the frictional properties of the tip–surface interface. The image shows the periodicity of the HOPG surface potential.

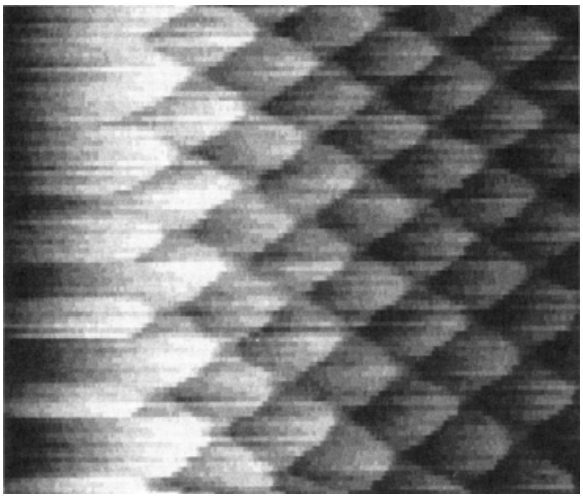


Figure 7. The frictional force in the x direction as a function of x and y. The intensity of the image is scaled to the friction force with the bright areas corresponding to 1.8×10^{-6} N. Only scans in the left-to-right direction are shown. The size of the image is 2 to 2 nm^2, and no correction from piezoelectric scanners was done. The load is 5.6×10^{-6} N, and the wire spring constant is 2.5 kN/m. Reprinted with permission from [67], C. M. Mate et al., *Phys. Rev. Lett.* 59, 1942 (1987). © 1987, American Physical Society.

The atomic scale stick and slip is believed to be responsible for the energy dissipation at the tip–surface interface. In such a case, a mechanism introduced by Tomlinson in 1929 [68] can be used to explain the energy dissipation. He proposed that some irreversible process in the passage of one atom past another must exist to explain the energy dissipation in pure conservative potentials.

The stick–slip movement of the scanning tip can be simulated in either one or two dimensions. Usually, a simple oscillator model as proposed by Tomlinson is used to simulate the FFM images in which the microscope tip slides over the corrugation of an atomically flat surface. This has been mainly performed in periodic potentials, generally described by sinusoidal potentials [69]. The calculations show that the FFM images present the surface potential periodicity instead of atomic resolution [70] and that the cantilever stiffness and scanning direction are important to the friction force images [71, 72]. The movement of the microscope tip can be understood by an analysis of the system's potential energy [6, Chap. 4]. The tip sticks in a surface minimum, until the elastic force in the tip–surface system overcomes friction. Then the tip jumps directly to the next surface minimum, leading to an atomically periodic stick–slip movement, as represented in Figure 8. As a consequence, only part of the surface potential is probed by the microscope tip, leading to images that contain only the periodicity of the analyzed surface. Recently, it has been shown that the FFM resolution is governed by the quotient between the average potential interaction energy and the average elastic energy stored on the tip before the jumps. It was also shown that there is an optimal velocity with which the scanning tip better senses the surface potential [74].

Studies on the influence of the normal force on friction revealed that at low normal forces, in contrary to what is observed in the macroscopic scale, the friction forces strongly depend on the real tip–surface contact area [75]. Friction is observed to be proportional to the contact area,

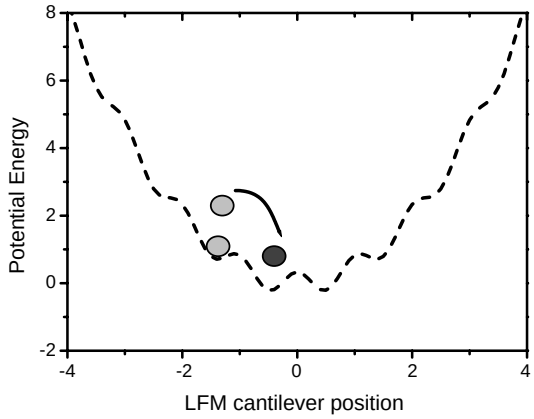


Figure 8. The system's total potential energy as a function of the LFM cantilever position. The dashed and solid lines represent the potential, while the gray and black dots represent the microscope tip position. Note that, as the cantilever is scanned, its tip gets pinned in a local surface minima (dashed line) until the elastic energy is high enough to overcome the friction barrier (solid line). Thus, the tip slides to the next surface minima (gray → black balls).

and the proportionality coefficient is defined as the interface shear strength. It has been shown by Kaneko et al. [76] that, if the microscope tip does a multiasperity contact with a surface, the dependence of friction with the normal force becomes linear because the contact asperities can increase with an increase in the normal force. They have also shown that if the tip does a single asperity contact with the surface, the relations between the friction and the normal forces are no longer linear. A change from multi-asperity to single asperity contact is shown to be caused by the capillary condensation of water on the microscope tip.

The influence of the environmental conditions on the friction experiments is not restricted to changes in the kind of tip sample contact. The role of water condensation was investigated in detail for silicon wafers, and the results for the friction forces revealed the importance of the surface energy in defining the thickness of the water film and its effects on the friction forces [77]. It was shown by Riedo et al. [78] that the capillary condensation of water strongly influences the dependence of friction on the scanning velocity. The formation of a capillary bridge increases the adhesion between the tip and a surface, leading to an increase in the force necessary to move the tip with respect to the surface. As the scanning tip moves faster, the time that the system has to built up the capillary decreases, leading to a decrease in the force necessary to move the tip on the surface.

In fact, the dependence of friction on the scanning velocity may be explained with a combination of three different models. At very low scanning velocity ($v \to 0$), the energy dissipation is shown to be mainly caused by the nonlinear dynamics of the moving parts [79–81]. At higher velocities, friction is shown to present a logarithmic dependence with velocity. This logarithmic dependence is explained by a combination of thermally activated stick–slip behavior [82, 83] and the kinetics of nucleation of water bridges between the moving parts [78]. In force microscopy the main energy dissipation process that is believed to occur when the tip is scanning on a surface is phononic excitation. Whereas at very low scanning velocities, the nonlinearity of the tip motion together with the coupling of its 2 degrees of freedom, is responsible for the energy exchange between its vibration modes [84, Chaps. 1, 6, and 7], at higher velocities, the discontinuous nature of the sliding tip movement is able to excite the normal modes of vibration on the scanned surface, raising its temperature.

4. NANOTRIBOLOGY RESULTS

4.1. Adhesion

When two surfaces in vacuum or in a gaseous environment come close, they usually experience attractive forces, such as Van der Waals forces. Once in contact, they are inevitably deformed because of their finite elasticity. Contact between two solid bodies can be described by the model of Johnson, Kendall, and Roberts (JKR) [85] or by the model of Derjaguin, Muller, and Toporov (DMT) [86]. Both models are based on an earlier analysis by Hertz [87], who considered two elastic bodies in contact under an external load but ignored attractive forces. In the JKR approach, the effective steady-state pressure in the contact area is assumed to be the superposition of the elastic pressure and of the attractive surface forces, which acts only over the contact area. As a result, a tensile force is necessary to separate the adhering surfaces. For two spherical particles this pull-off force is given by

$$F_{\mathrm{JKR}} = 3\pi R\gamma \tag{1}$$

where γ is the effective solid surface energy and R is the reduced radius of curvature of the two surfaces ($R = R_1 + R_2/R_1R_2$).

The DMT model is an alternative model that also accounts for noncontact forces in the vicinity of the contact area. It predicts a slightly higher pull-off force of

$$F_{\mathrm{DMT}} = 4\pi R\gamma \tag{2}$$

Both models are limiting cases of a more general description [88]. The JKR model is appropriate for large, soft bodies with high surface energy, and for hard solid bodies with low surface energy, the DMT model should be applied [89, 90].

For an AFM tip interacting with a carbon film surface, we can model the tip–surface interaction as a sphere contacting a flat surface, the DMT theory describes the adhesion force as

$$F_{\mathrm{ad}} = 2\pi R W_{\mathrm{ad}} \tag{3}$$

where R is the radius of curvature of the sphere and W_{ad} is the work of adhesion defined as

$$W_{\mathrm{ad}} = 2(\gamma_{\mathrm{tip}}\gamma_{\mathrm{surface}})^{1/2} \tag{4}$$

where γ is the surface free energy of the tip and of the surface [91]. From this model it is seen that adhesion forces scale directly with the radius of curvature of the tip. Therefore, to correct for changes in tip geometry and to probe the inherent properties of the carbon film surface, one needs to normalize the adhesion forces by the radius of curvature of the AFM tip [92].

When two surfaces come in contact, contact occurs on a large number of asperities. To develop a fundamental understanding of friction and wear mechanisms in lubricated contacts, experiments in a single asperity contact need to be conducted. The development of the atomic force makes possible the investigation of the adhesion forces and its influence on the tribological performance of coatings in a single asperity contact regime.

One of the first discrepancies between the macroscopic behavior and the friction on the nanometer scale was revealed by studies designed to investigate the dependence of friction forces with the contact area. The fundamental friction law, first proposed by da Vinci and reproposed by Amonton more than three centuries ago, can be written as

$$F_F = \mu F_L \tag{5}$$

where F_F is the friction force, μ is the friction coefficient, and F_L is the normal load force. The friction coefficient is constant for a given pair of materials in a wide range of

loads and independent of the apparent contact area. This condition is fulfilled when the contact area, A, is proportional to the normal load. It was shown that for rough surfaces (multiasperity contact) in an elastic regime, it occurs if the asperities follow a statistical height distribution [93–95]. Moreover, when the normal load increases and plastic deformation occurs, an increase of the contact area is verified so that the relation contact area, $A \sim F_L$, remains valid and so follows Amonton's law.

In a single-contact regime the situation is different. In fact, load-dependent studies of frictional properties of graphite, diamond, amorphous carbon, and fullerene thin films were performed in argon atmospheres and ambient conditions by Wiesendanger and collaborators [96] using specially prepared silicon tips with well-controlled tip point shape and radius. Their results reveal that in the low load regime, that is, in a regime where wear and plastic deformations can be neglected, the friction force did not follow direct dependence with the applied load. The data were explained using a theoretical model based on a Hertzian-type tip–surface contact, which predicts a dependence of the contact area with the normal force of $A \sim F_L^{2/3}$ [87]. With this result, the authors show that friction is proportional to the contact area and that the shear stress was constant within the applied pressure, leading to a $F_F \sim F_L^{2/3}$ dependence of the measured frictional forces. These experiences showed that single asperity contact mechanical models are valid for tip radii down to a few nanometers, at least when the tip has a well-defined tip–surface contact area. In nanoscale the normal force must take into account the attractive adhesion forces

$$F_N = F_L + F_a \tag{6}$$

where F_N is an effective normal force that takes into consideration both the normal load F_L and the adhesion force F_a. The adhesion force is taken as the maximum negative bending force of the cantilever on the tip approach curve.

Many experimental works are devoted to determination of the adhesion forces in carbon film surfaces and many other surfaces [92]. In a pioneer investigation, Binggeli and Mate [97] showed that the adhesion force measured in different ambient, relative humidities ranging from 80 to 98%, can be considered as a constant for a-C films deposited by sputtering.

The effects of hydrogen on adhesion force of a-C:H films were carefully investigated by Perry and collaborators [98]. They showed that the adhesion force increases with the hydrogen content in films deposited by sputtering in Ar–H_2 atmospheres. Because they used a tip with the same chemical composition, oxidized tungsten in this case, they concluded that the observed increase arises from an increase in the surface free energy of carbon films that follows the hydrogen content in the films. The friction coefficient measured in the same samples also increased with the hydrogen content, following the adhesion force behavior.

The effect of nitrogen incorporation was also investigated for films deposited by ion beam sputtering using a second argon beam as an assistance beam [55]. The authors, using an AFM with a Si_3N_4 tip, operating in air, showed an increase from 25 to 32 nN for adhesion forces measured in a-C and a-CN_x films, respectively. Despite higher adhesion at the interface, a-CN_x films showed a smaller friction coefficient than did a-C films—0.14 and 0.28, respectively. The contact radius calculated using the Hertz model was about 0.2 nm and almost equal for both coatings. The authors concluded that the shear strength of CN_x was lower than that of carbon and could partially explain why friction is lower in CN_x than in carbon films.

4.2. Friction

As previously mentioned, the laws that control the friction at the macroscopic scale are well known: the friction force (F_F) increases with the applied normal load (F_L) and it is independent of both the contact area and the velocity [66]. The studies carried out since the invention of the friction force microscope reveals that some of these laws are no longer valid on the nanometer scale. To investigate this new experimental domain, the study of carbon-based materials occupies a special place. A recent review of tribological properties of carbon films shows that, in all environments, the macroscopic tribological behavior of these films was controlled by an interfacial graphitic transfer layer formed during friction [99]. However, in nanoscale, those energy-dissipating mechanisms are far from being well understood. Some authors claimed that the dissipation of friction energy could take place by phonon excitation [100]. A second proposed mechanism takes into consideration the dissipation via electron–hole pair excitation [101]. Amorphous carbon films represent a remarkable opportunity to investigate friction mechanisms because materials with different microstructures, the sp^2/sp^3 ratio, for example, can be easily synthesized.

The FFM experiences showed that single asperity contact mechanical models are valid for tip radii down to a few nanometers, but only when the tip has a well-defined tip–surface contact area. These experimental results also show that, in nanoscale, the normal force must take into account the attractive adhesion forces; that is, $F_N = F_L + F_a$. Wiesendanger and collaborators [96] show that, for measurements performed in argon atmospheres, the friction in C_{60} films is higher than those for a-C films and diamond surfaces. In ambient air, the friction is higher than friction forces measured in a water-free ambient. For atmospheric exposed surfaces, a-C films have friction higher than that for diamond. In both cases, the lowest friction forces were always determined for HOGP samples.

Following this first work, several other groups performed a comparative study of the friction properties of different carbon-based films. Among them we mention the investigation performed by Riedo and collaborators [102]. They studied diamond-like films with different fractions of sp^3-hybridized carbon atoms and CN_x films with different contents of nitrogen ($0.2 < x < 0.3$), both deposited by laser ablation. The Raman results obtained from the samples show that, depending on the laser fluence, the content of sp^3 hybridization can be controlled, and it is higher for films deposited at the higher power laser in a nitrogen-free atmosphere. These films have 53% sp^3-hybridized carbon atoms and are called diamond-like carbon (DLC). The FFM measurements were performed in ambient conditions using a

silicon nitride tip. To avoid plastic deformation, the loads were restricted to 20 nN.

As was observed before by Bushan and Sundarajan [103, 104], the topographical images of CN_x and a-C films obtained simultaneously with the friction force images revealed that the friction forces map essentially follows the surface topography, as can be seen in Figure 9. In this figure the topographical image of a CN_x film was compared with the friction image in the forward and backward scan direction taken simultaneously. In Figure 10, we show the friction forces for HOGP, CN_x, and DLC films for an applied load of 12 nN. The figure shows the forward and backward scans with the friction force being proportional to the difference between the two. It is clear from the figure that the friction force in HOPG is by far the lowest one and that the friction force measured on the CN_x film is of the same order of magnitude as the one obtained from a DLC, but slightly lower. The authors claimed that the friction forces obtained from a-C films with sp^3 fractions in the range of 34–53% are equal within experimental errors. Similar behavior was observed for CN_x films with different nitrogen content; that is, in this case the friction forces are independent of the amount of nitrogen incorporated in the films.

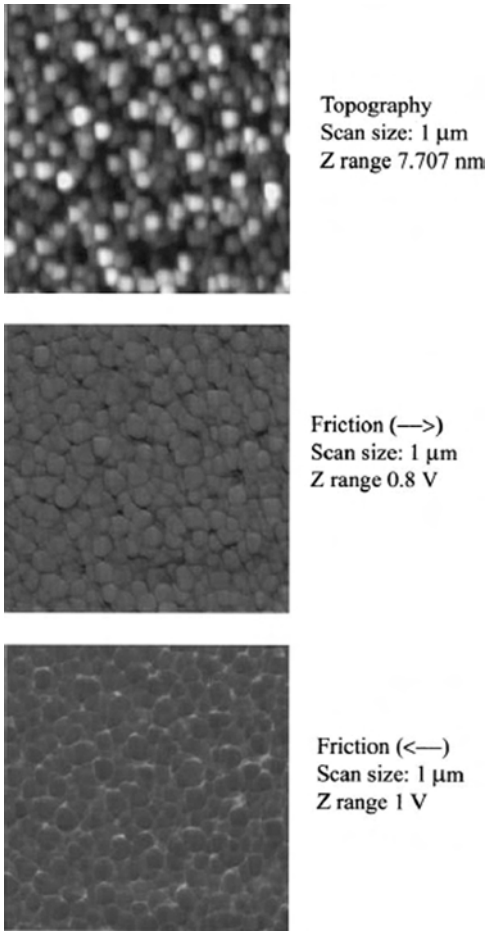


Figure 9. Topographical image (top) of a CN_x film taken simultaneously with a friction image in the forward (→) and the backward (←) scan direction (middle and bottom image, respectively). Reprinted with permission from [102], E. Riedo et al., *Surf. Sci.* 477, 25 (2001). © 2001, Elsevier Science.

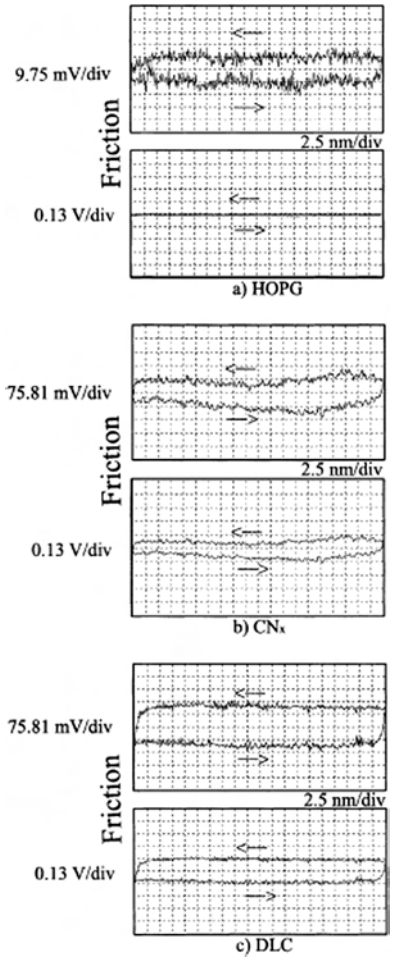


Figure 10. Friction forces for the samples of (a) graphite (HOPG), (b) CN_x, and (c) DLC. Each graph presents a scan width of 50 nm and two force signals corresponding to foreword and backward scans. Reprinted with permission from [102], E. Riedo et al., *Surf. Sci.* 477, 25 (2001). © 2001, Elsevier Science.

The investigation of the friction force as a function of the normal load reveals that CN_x film results show linear dependence between the friction force and the load force, indicating a multiasperity contact, while the results obtained from a-C films nicely follow a two-thirds power law, $F_F \sim F_N^{2/3}$, as expected from a single asperity contact. Because a similar silicon nitride tip is used in all instances, the differences observed between DLC and a-CN_x films can be attributed to the lower surface roughness of the former. Because on nanoscale the Amonton's law $F_F = \mu F_L$ is no longer valid, the friction coefficient defined in that way is not well suited for comparing tribological behavior of different materials in the case of single asperity contact. The friction coefficient was determined by the linear regression of F_F versus F_N, that is, taking into account the adhesion forces, and the following results were obtained: $\mu = 0.009 \pm 0.003$ for HOPG, $\mu = 0.167 \pm 0.03$ for CN_x films, and $\mu = 0.208 \pm 0.03$ for DLC films. These results seems to support the phonon mechanism as the dominant one for energy dissipation, because the friction coefficient in DLC films is independent of the sp^3 fraction films, and, more important, the DLC films have

a higher friction coefficient compared with HOPG. If an electron–hole excitation mechanism played an important role, the opposite behavior should be expected.

Thin CN_x films, 1–10 nm thick, deposited by ion beam-assisted deposition (IBAD) were investigated by friction force microscopy [105]. Raman spectra from these films are quite similar to those obtained from films deposited by laser ablation and mentioned previously, suggesting the same microstructure [102]. Forward–backward scan cycles were obtained and the friction coefficient determined in the wear-less regime (load of 1 μN) is on the order of 0.12, a value that increases to 0.22–0.24 for higher loads (10 μN). No thickness dependence was observed in the range of 1–10 nm. These values for the friction coefficient were on the same order as the previous ones. Khurshudov et al. [55] using an AFM equipped with a silicon nitride tip (10–20 nm radius) also measured the friction forces for a-CN_x ($x = 10$ at.%) films deposited by IBAD. The measurements were performed at a velocity of 10 μm/s and at ~60% of relative humidity and show that the friction force measured for a-CN_x films was systematically 20% lower when compared with the friction force measured in a-C films deposited by the same technique, confirming the general trend showing that the friction force (or friction coefficient) measured in CN films is slightly smaller compared with the values obtained from a-C films deposited with the same technique using similar deposition conditions.

Films with different contents of nitrogen deposited by PECVD, a-C:H(N) films, were also investigated by friction force microscopy using a silicon nitride tip with measurements performed in air [106]. The results are summarized in Figure 11. It is clear from this figure that, despite a substantial increase of the sp^2 fraction upon nitrogen incorporation (data taken from Ref. 107), the friction coefficient is constant, supporting the idea that phonon excitation is the dominant energy dissipation mechanism. The increase in film surface roughness with the increase of the sp^2 fraction was observed before for a-C films deposited by C^+ beams [108] and seems to play a secondary role in the definition of friction forces in this case.

A comparative FFM study of a-C films deposited by different techniques was carried out by Sandararajan and Bushan [109]. Films deposited by electron cyclotron resistance (ERC)–chemical vapor deposition (CVD) using ethylene as a precursor gas, rf sputtering of a graphite target in argon atmosphere, FCVA, and direct ion beam deposition from methane dissociation were studied. The results for the friction coefficient were obtained with a diamond tip at a 4 μm × 4μm scan size in a load regime where no wear was observed. The results for friction coefficients were determined for films with thickness in the range of 3.5–20 nm. The thinner films have higher friction coefficients independent of those deposited using the same technique. These results contradict the results obtained by Bai et al. [105] for CN_x films mentioned above. The systematic investigation of the coating failure by the AFM topography image reveals the direct correspondence between high friction and coating failure area. This problem is critical in the case of sputtered carbons films, where reliable coatings with thickness below 5 nm could not be obtained.

The fundamental laws of tribology predict that the friction is independent of the velocity. Friction force microscopy measurements on carbon-based compounds have also been performed to investigate this question on the nanometer scale. In a first report, Zwörner et al. [110] determined that the friction forces are constant over a wide range of velocities. Metal-containing carbon films, Me–C:H, with Me = Au and W were also investigated, and the results show that in the range of 0.1–10 μm/s, the friction force measured by the FFM is constant [111]. In another investigation, constant friction coefficients were reported for ta-C and FLC films in the velocity range of 0.1–40 μm/s [112]. The results are shown in Figure 12. In this case, the ta-C films were deposited by FCVA while the fullerene-like carbon films were deposited using an unfiltered LHPA discharge,

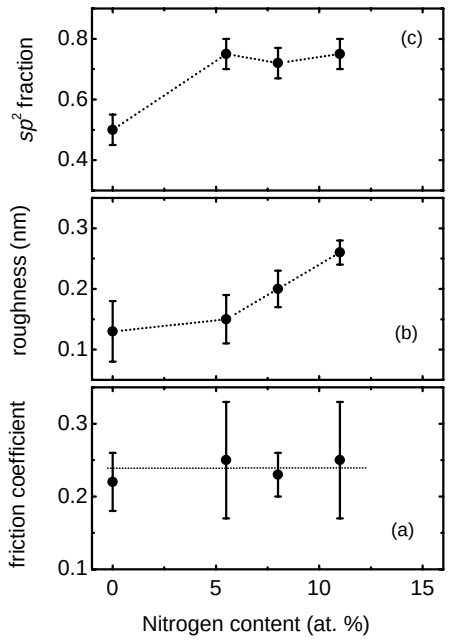


Figure 11. Fraction of carbon atoms in sp^2 hybridization (a), surface roughness (b), and friction coefficient (c) for a-C:H:N films deposited by PECVD. Reprinted with permission from [106], R. Prioli et al., *J. Vac. Sci. Technol., A* 14, 2351 (1996). © 1996, American Vacuum Society.

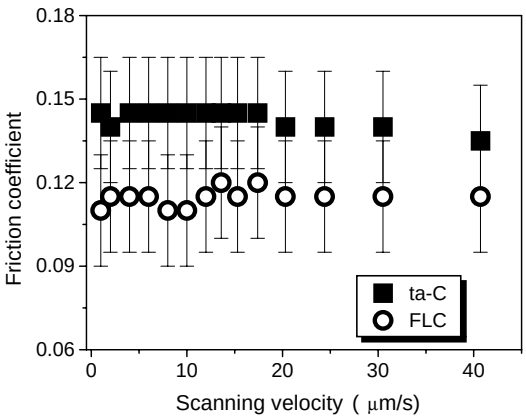


Figure 12. Friction coefficient as a function of the scanning velocity for ta-C and FLC films [102]. The lines are only to guide the eyes.

introducing the carrier gas (nitrogen) via a 1 mm hole in the graphite (99.99% purity) target. The friction force measurements were carried out using an AFM operated in the friction force regime. The friction force (forward and reverse cycles) images were acquired at constant normal load, and the total normal force is composed of the adhesion force between tip and sample surface and the applied load. The cantilever-bending constant and the calibration method proposed by Liu et al. [66, 113] were used to obtain the absolute values of the friction coefficients. Nanostructured carbon films deposited by supersonic cluster beam were also studied by friction force microscopy [114]. These clusters have a fullerene structure and linear and planar structures [115]. The friction coefficients reported are in excellent agreement with the values obtained for FLC films [112]—$\mu = 0.12 \pm 0.2$—considering that the values reported for the nanostructured films depend on the region of the sample and the size of the clusters and varies from 0.10 to 0.14, the higher value being obtained from films deposited with the larger clusters.

Gnecco et al. [83] determined a logarithmic increase of the friction force with the velocity for ultrahigh vacuum friction force microscopy measurements performed on NaCl (100) single crystals. These results were interpreted in terms of a thermally activated stick and slip behavior and the friction being due to the cohesive forces between the two surfaces in contact. It was shown that these forces increase with the scanning velocity and this logarithmic dependence was explained by a modified one-dimensional Tomlinson model [68]. Similar velocity dependence, that is, a logarithmic increase of the friction forces within the velocity was observed for DLC coated silicon substrates for friction force microscopy measurements performed on air [116].

A recent investigation carried out by Riedo et al. [78] showed that the presence of water in the ambient air changes this situation. In fact, they investigated the friction dependence of amorphous carbon films with different surface energy and showed that the friction forces were velocity dependent, confirming the logarithmic character of this dependence. However, not only increasing behavior was observed. For other materials, such as CrN films, a decrease in the friction forces with velocity was verified. This fact occurs when the measurement is performed in air and the dependence with the velocity becomes stronger for atmospheres with higher humidity. In this situation, capillary effects due to water vapor condensation must be taken into consideration. In a humid environment, the force resulting from the condensation capillary liquid bridges at many different points along the tip–sample contact area is responsible for a velocity dependence force. When the tip is at a fixed position, this capillary force increases as a function of time, because of the continuous formation of liquid bridges at the tip–sample interface. When the tip is scanning the sample surface, the time available for liquid bridges decreases as a function of the velocity, and so, the capillary forces. Hence the capillary forces acts in opposition to atomic stick and slip, causing a decrease of friction with increasing velocity. This effect obviously depends on the hydrophobicity of the sample surface. For hydrophilic surfaces, this mechanism can be predominant and friction decreases with velocity. The velocity dependence behavior observed experimentally can be understood as a competition between two mechanisms, the cohesive forces responsible for the stick and slip movement and the kinetics nucleation of water meniscus between the moving parts.

Experimental data obtained in air from a-C:H and a-C:H:F films deposited by PECVD are shown in Figure 13 [117]. The results clearly show the existence of the logarithmic velocity dependence. It is observed that the increase of the hydrophobicity of the carbon–fluorine films, represented in this case by the contact angle, is responsible for the progressive increase of the slope of the friction forces dependence with velocity, in agreement with the model proposed by Reido et al. [78].

The dependence of the friction coefficient with the surface energy is illustrated in Figure 14, where the friction coefficient obtained from hydrogenated carbon films deposited by sputtering [98] and a-C:H and a-C:H:F films deposited by PECVD are plotted as a function of the contact angle [97]. The friction coefficients were determined by the linear regression of F_F versus F_N, that is, taking into account the adhesion forces, and show a clear correlation with the surface energy: the higher the contact angle, the lower the friction.

These results for the friction coefficient show the importance of the study of the adhesion forces and capillary condensation for the complete description of the friction behavior. The first systematic investigation of this problem was carried out by Binggeli and Mate [97, 118]. In Figure 15 their results for both the friction coefficient and adhesion force as functions of the relative humidity are presented for a-C films and silicon dioxide. The sliding speeds were on the order of micrometers per second. The slope of the friction versus load curve provided the friction coefficient and the extrapolation to zero loads, the adhesion force. In this experiment, the linear dependence between friction and applied load was an indication of the multiasperity contact among the AFM tip and sample surface. The results show that the friction coefficient decreases for silicon oxide surface but not for carbon films upon the increase of the relative humidity. The same behavior was observed for the adhesion force. The difference between the two samples was explained in

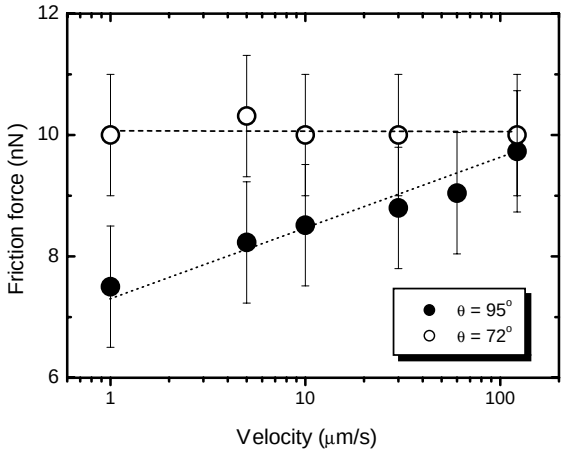


Figure 13. Friction coefficient as a function of the scanning velocity for films with different contact angle: $\theta = 95°$ for a-C:H film and $\theta = 72°$ for a-C:H:F film [106]. The lines are only to guide the eyes.

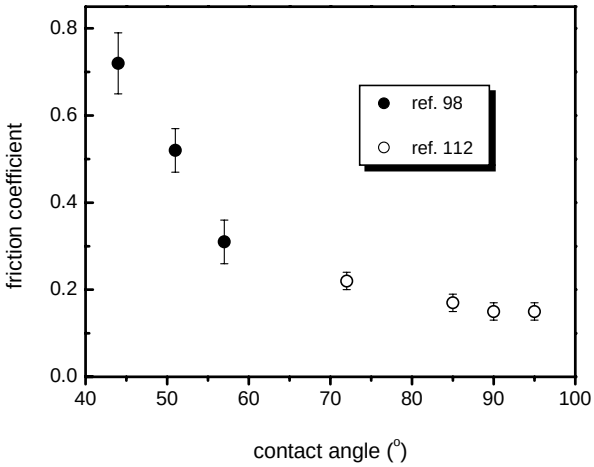


Figure 14. Friction coefficient as function of the contact angle for a-C:H films deposited by sputtering [108] and a-C:H and a-C:H:F films deposited by PECVD [106].

terms of the hydrophilic nature of the oxide, where water molecules are strongly bounded to hydroxyl groups. The friction coefficient for a-C films is nearly constant and equal to 0.3. A similar experiment was performed for a lower relative humidity regime (0–60%) and showed a slight increase of the adhesion force upon a humidity increase [92]. This apparent contradictory behavior can be explained as follows. For very low relative humidity, there is an increase in the number of the meniscus bridges, which results in a progressive increase of the friction. It occurs more or less independently on the surface energy of the samples. When the relative humidity is higher, the hydrophobicity of the film surface plays the decisive role.

The influence of hydrogen content in a-C:H films tribology is important in either the macroscopic or nanoscopic scale. In fact, in a review article by Erdemir [119] the role of hydrogen in the tribological behavior of diamond and diamond-like films was extensively discussed. This article showed that, for films deposited by PECVD in CH_4–H_2 atmospheres, the friction coefficient measured by conventional techniques in dry nitrogen (humidity = 0%) can be extremely low ($\mu = 0.003$) when the samples were grown in a very high H/C ratio in the precursor atmosphere (H/C = 10). On the other hand, it can be very high ($\mu = 0.65$) when measured in a hydrogen-free film. With an increasing the humidity, the friction coefficient of the hydrogen-free film dropped to 0.25 and for the film deposited in a hydrogen-rich atmosphere the friction coefficient increased to 0.06. The superlubricity of the a-C:H films was explained in terms of the almost complete elimination of covalent bonds at the carbon film surface by hydrogen passivation of carbon dangling bonds [119, 120]. The C—H bonds (stronger than C—C bonds) inhibit the covalent interaction with the atoms of the counterface materials and, thus, cause low adhesion and friction.

These arguments must be valid on the nanometer scale. In Figure 16 the results obtained by friction force microscopy from a-C:H films deposited in $CH_4 + H_2$ atmospheres by ECR–CVD [121] are shown. The friction coefficients obtained by Erdemir and collaborators [122] are also presented. The trend observed for both sets of data are the same: a decrease in the friction coefficient for films deposited in a hydrogen-rich atmosphere. It is interesting to note that films with high hydrogen content (up to 40 at.%) deposited by sputtering in Ar–H_2 atmospheres [98, 123] show exactly the opposite behavior; that is, the friction coefficient increases with the hydrogen content of the films. The different effects of the plasma film surface interaction in the case of $Ar + H_2$ mixtures in a sputtering chamber and in the case of hydrocarbon–H_2 mixtures in PECVD deposition are still not well understood. Probably they are responsible for the behavior observed for the a-C:H films, because the dangling bond density in the uppermost surface layer of a-C:H films certainly depends on the deposition technique used.

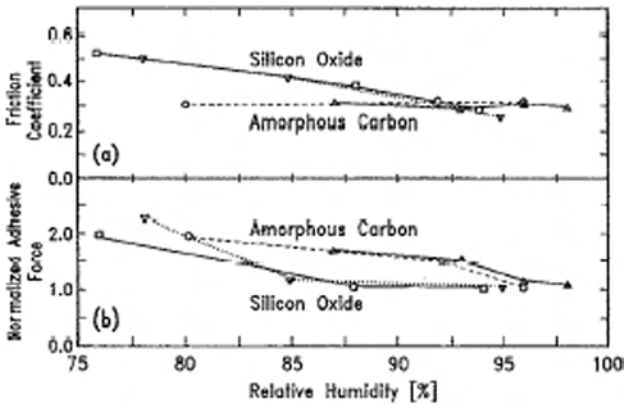


Figure 15. Friction coefficient and normalized adhesive forces during sliding as a function of relative humidity. Two typical measurement series for clean silicon wafers with silicon oxide and amorphous carbon are shown. Adhesive forces have been normalized by setting the lowest value to 1 to account for different tip radii in the separate experiments. Reprinted with permission from [97], M. Bingelli and C. M. Mate, *Appl. Phys. Lett.* 65, 415 (1994). © 1994, American Institute of Physics.

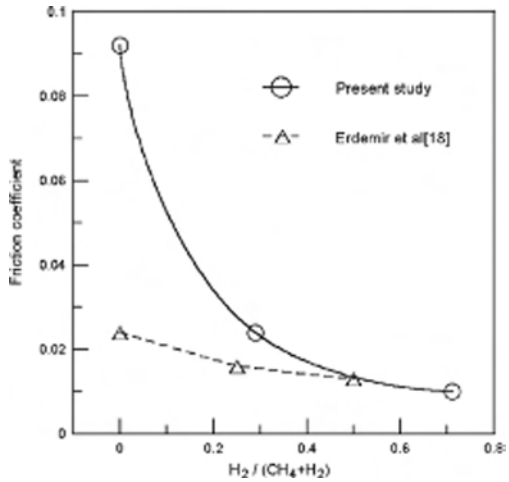


Figure 16. Friction coefficient as a function of the hydrogen content in CH_4-H_2 atmospheres. Reprinted with permission from [121], T.-H. Fang et al., *Thin Solid Films* 396, 166 (2001). © 2001, Elsevier Science.

4.3. Wear

The demands resulting from the continuous increase of the storage density of computer hard disks and from the reduction of the loads and dimensions used in microelectromechanical systems; that is, thinner protective and anticorrosive coatings with lower surface roughness, together with the spreading of the use of the AFM motivate a large experimental effort to understand the wear behavior, in nanoscale, of carbon-based films deposited by different techniques [124]. In fact, it was demonstrated that it is possible to enter in a subnanometer scratching resistance testing regime using a standard AFM with diamond-tipped cantilevers [103]. As was also discussed before for the friction studies, a comparison of results obtained by different groups is not straightforward because the authors frequently did not report a complete structural characterization of the films under investigation. The comparison of the absolute values of wear data, the scratch depth, for example, is nontrivial, because sometimes not taken into account is the fact that the constant applied force on the cantilever cannot overcome the extra resistance force arising from the larger tip–surface contact area at deeper groove depths. On the other hand, some authors explicitly take into consideration the blunting of the tip when presenting their data [109].

Hydrogenated carbon films are a good example of this situation. These films can be deposited by sputtering with different hydrogen content by changing the ratio between Ar and H_2 gases in the camera. The control of the hydrogen content can also be achieved in a-C:H films deposited by PECVD techniques by changing the hydrocarbon–hydrogen precursor gas mixtures. Hydrogenated films deposited by sputtering were investigated by Jiang et al. [123]. The 25 nm thick sputtered films were scratched with an AFM diamond tip with a loading force of 28 μN and were imaged by the same tip after 16 wear cycles. The results show that the scratch depth increases from 4 to 30 nm with the increase of hydrogen content in the films, which varies from 2 to 40 at.%. The scratch depth of 30 nm measured for the hydrogen richest film means that the film is completely removed by the AFM tip. Nanoindentation measurements show that the hardness decreases with the increase in the hydrogen content from 9 to 6 GPa. A similar investigation was carried out by Wiens et al. [125]. From similar tests, they show that the scratching resistance decreases with the amount of hydrogen incorporated in the film for loads of 17 and 23 μN.

Sputtered hydrogenated carbon films were also studied by Anoikin et al. [126]. They deposited 10 nm thick a-C:H films onto AlMg/NiP/Cr/Co-based magnetic layer smooth disks (average surface roughness of 0.5 nm). The tip velocity was 1 μm/s and the load force was in the range of 10–15 μN. The films were deposited in Ar (a-C films), $Ar + H_2$ (a-C:H films), and $Ar + N_2$ (a-C:N films), and the amount of hydrogen and nitrogen was measured and Raman analysis was performed on each sample. Despite the fact that it is not possible to obtain the sp^2/sp^3 ratio directly from the Raman results, the position of the G-band can be used, in a qualitative way, to infer an increase (shift to higher wavenumbers) of the amount of sp^2-hybridized carbon atoms in the film or of the size of the graphitic domains [127]. The results are resumed in Table 1. They show that for films with nearly the same amount of hydrogen, the scratch depth decreases when the sp^2 fraction increases. When we compare the results for scratch depth obtained for films deposited by sputtering with those obtained by ion beam deposition of hydrocarbon species, in the same experimental conditions, this correlation is even clearer. The results obtained from sputtered films and presented in Table 1 show a direct correlation between hardness and the scratch depth, considering that films with higher sp^3 fractions usually have higher hardness [15, 63].

Table 1. Hydrogen content, G-band position, and scratch depth for several a-C films [123].

Materials	H content (at.%)	G-band position (cm^{-1})	Scratch depth (nm)
a-C	13.5	1582	14
a-C:H	25.4	1575	7
a-C:N	9.1	1569	4
a-C (IBAD)	18.9	1535	0.5
Graphite [127]	—	1584	—

Hydrogenated carbon films deposited by ECR–microwave plasma CVD were also studied for different chemical composition [121]. In this case the precursor atmosphere was a mixture of H_2 and methane, and the authors varied the amount of hydrogen in the mixture from 0 to 71%. The wear depth as a function of the number of wear cycles is presented in Figure 17. A load of 100 μN was applied, and each sample was submitted to 20 wear cycles. The wear rate decreases with the hydrogen content of the films. In this work, Raman results were presented and clearly show that a-C:H films have an increasingly graphitic character when the gas precursor mixture became richer in hydrogen: the G-band shits to higher wavenumber and its width reduces upon the increase of the amount of H_2 in the plasma atmosphere. The Raman results suggest a hardness reduction upon the hydrogen incorporation. The same group investigates the effect of the self-bias voltage in films deposited by ECR–microwave plasma CVD using a CH_4 (50%)–H_2 (50%) mixture [128]. Again the Raman result revealed an

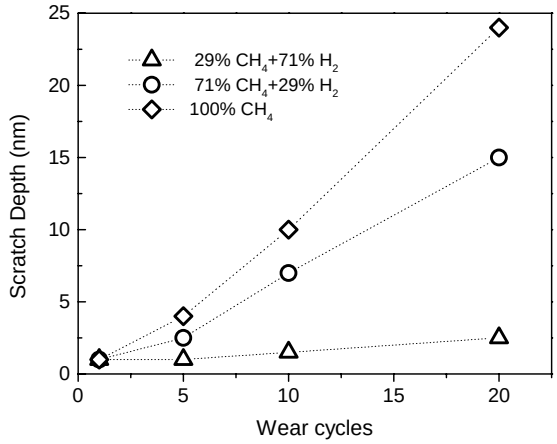


Figure 17. Wear depth as a function of the number of cycles for different hydrogen content in CH_4-H_2 atmospheres. Reprinted with permission from [121], T.-H. Fang et al., *Thin Solid Films* 396, 166 (2001). © 2001, Elsevier Science.

increase in the graphitic character of the films upon the self-bias increase. The wear test performed under a load of 200 μN, revealed the better wear resistance performance for the a-C:H film deposited at the higher bias voltage. Because the amount of hydrogen effectively incorporated in the films is not reported in this article, as well as the film hardness, it is difficult to compare these results with other data reported in the literature that shows an opposite behavior, especially in the case of the hydrogen content of the films.

A systematic investigation of the wear resistance for a-C films deposited by different techniques was carried out by Sundararajan and Bhushan [109]. Films deposited by ECR-CVD, sputtering, FCVA, and direct ion beam deposition were studied. They performed the scratch test using a pyramidal diamond tip mounted onto a platinum-coated stainless steel of stiffness of 40 N/m. The test consisted of 10 cycles over a scan length of 5 μm with a velocity of 10 μm/s and at a given load. In Table 2 we compare the data obtained at a load of 60 μN for films with the same thickness, 20 nm. Despite the lack of structural characterization of the tested samples, we can expect, as also claimed by the authors, that the film hardness closely follows the sp^3-hybridized carbon atom fraction. As previously discussed the results listed in Table 2 revealed that the scratch depth increases when the film hardness decreases.

Sundararajan and Bhushan [109] also performed wear tests. In this experiment, the tip scans a 2 μm $\times$ 2 μm area at a required load at 4 μm/s for a certain number of cycles. Then, the normal load was lowered to 1 μN, and the scan size was set to 4 μm $\times$ 4 μm to image the wear region. A detailed description of this procedure can be found in Ref. 129. Wear is not uniform, but it is initiated at the nanoscratches, indicating that surface defects (with high surface energy) act as initiation sites [103]. The wear test results showed that films 20 nm thick deposited by FCVA and ECR–CVD have excellent wear resistance for loads up to 80 μN. These deposition techniques provide films with similar performance even for thickness as low as 5 nm, but in this case for a maximum load of 20 μN. For very thin films, 3.5 nm thick, ECR–CVD deposited films have better performance. In fact, a-C films 3.5 nm thick deposited by FCVA are easily delaminated. On the other hand, 20 nm thick films deposited by ion beam deposition show a negligible wear depth for tests at 60 μN. For higher loads, the film is completely removed by the AFM tip. However, thinner ion beam deposited films have better wear behavior than the films deposited by a filtered cathodic arc and a performance quite similar to the one obtained by films deposited by ECR–CVD. In fact, for films 5 nm thick, the better performance was obtained for ion beam deposited films that show wear depths of around 2 nm for loads of 25 μN. At this load the AFM tip completely removed the films deposited by the other techniques. For thickness of 3.5 nm, both ECR–CVD and ion beam deposited films have the same performance under loads of 20 μN. On the other side, sputtered carbon films have the worst wear resistance, independent of the thickness. The results obtained for thicker films show that at least for a-C thick films the wear depth, as well as the scratch depth, closely follows the film hardness. However, for thinner films different behavior was observed. The better wear performance was obtained for ion beam and ECR–CVD deposited films that have an elastic modulus similar to that of the silicon substrates, whereas films deposited by a cathodic arc have the higher elastic modulus. This fact suggests that thinner coatings share load with the substrate, causing its deformation. Then, cracks can be created and propagate, resulting in the film delamination. The high internal stress of the cathodic arc deposited coatings aids propagation of cracks and facilitates spreading of the failed region compared with other coatings [130]. In this way, an analysis of the wear depth for very thin films must take into account the critical load required to fail the coating during the test.

Table 2. Hardness, elastic modulus, and scratch depth for carbon films prepared by different deposition techniques [1090000].

Material	Deposition technique	Hardness (GPa)	Elastic modulus (GPa)	Scratch depth (nm)
Diamond	—	100	1000	—
a-C	FCVA	24	240	$\sim$0
a-C:H	ECR-CVD	22	180	1
a-C:H	Ion beam	19	140	8
a-C:H	Sputtering	15	140	20

ta-C films were studied by Martínez et al. [131]. The films were deposited by filtered cathodic vacuum arc. The wear test was performed with an AFM equipped with a diamond tip that scans an area of 3 μm $\times$ 3 μm with a scanning frequency of 1 Hz. To obtain films with different hardness the substrate voltage was varied from -20 to -350 V. Films with thickness of around 100 nm were scratched with loads of 10 and 40 μN and the scratch depth can hardly be measured with the diamond tip and was estimated to be of the order of 1 nm, similar to those reported in Table 2 for films deposited by FCVA with 20 nm, and much lower than the 27 nm for bare silicon previously reported by Bushan and Koinkar [132]. These results confirm the high wear resistance of thick ta-C films and show a direct correlation between film hardness and wear resistance.

A scratch test can also be performed with very low loads and soft tips such as the silicon nitride tip to obtain information about the wear mechanism in hard carbon films. An AFM 3D image of a scratch profile on a ta-C film is shown in Figure 18 [112]. In this case 30 nm thick films were deposited by FCVA with a bias of -80 V applied to the substrate to produce films with an sp^3/sp^2 ratio of about 80%. The scratch test was performed using a Si_3N_4 tip of $k = 0.77$ N/m, scanning velocity of 2 μm/s, 1024 cycles, and a total normal force of 470 nN. The image was obtained with the same tip used to scratch the surface.

The scratch depth as a function of the normal load was plotted in Figure 19 for ta-C films. In the macroscopic scale, wear of ta-C films is strongly influenced by a shear layer of graphite that is built up at the interface of the slider tip and the ta-C surface during sliding. Although this graphitic layer may reduce the dissipated energy at the interface, it may also increase the surface wear. As the normal force is increased, the friction force also increases, leading to larger energy dissipation at the tip–surface interface [133]. However, on a nanometer scale, the presence of a low-density layer sp^2 surface layer, approximately 1 nm thick for ta-C

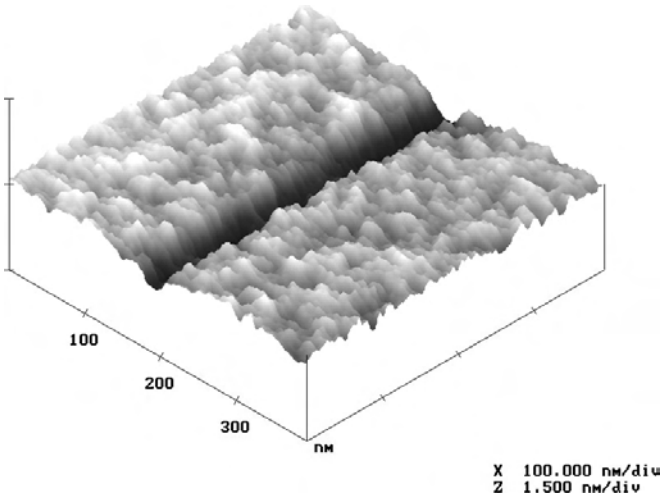


Figure 18. The result (topography image) of a scanning scratch test on a ta-C film performed with a Si_3N_4 atomic force microscope tip of $k = 0.77$ N/m, scanning velocity of 2 μm/s, 1024 cycles, and a total normal force of 470 nN [102].

films deposited with ion energy of ~100 eV, at the surface of the film due to the subplantation mechanism during film growth probably plays an important role [134]. This layer may be easily worn during the scanning process leading to the wear observed at the film surface. It is proposed that most of the wear observed at the ta-C films surface is caused by the removal of this low-density layer already present at the film surface. In fact, the measurement shows that after the complete removal of this layer the wear process stops and the wear depth remains constant, despite the increase of the normal force.

The influence of the number of scanning cycles on the ta-C wear can be seen in Figure 20. The normal force applied by the tip on the surface was kept constant at 470 nN and the humidity was 43%. The wear depth increases for scanning cycles up to 512 (forward + backward scans), above which it remains constant. This fact is in agreement with the hypothesis mentioned above that the microscope

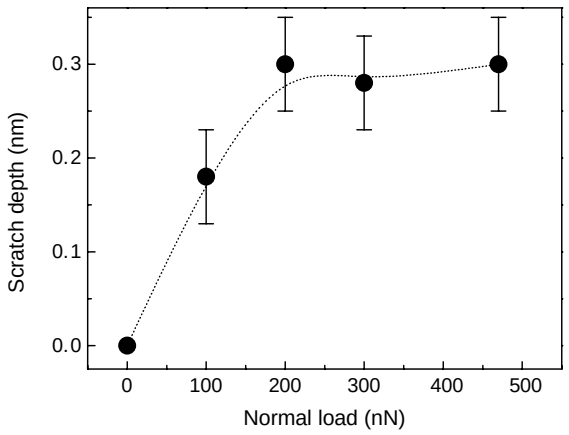


Figure 19. The scratch depth as a function of the normal load for ta-C films [102]. The line is only to guide the eyes.

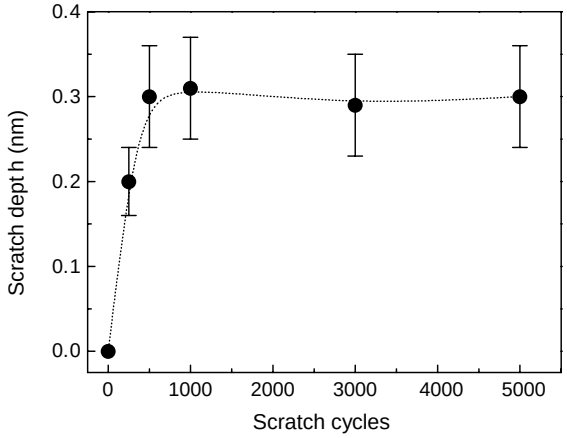


Figure 20. The influence of the number of scanning scratch cycles on the wear of ta-C films [102]. The line is only to guide the eyes.

tip only removes the low-density graphitic layer present at the ta-C film surface.

The effect of the film composition on the wear resistance has received increasing attention with the introduction of carbon–nitrogen films as protective coatings. One of the first efforts to establish a comparison between the scratch resistance, in nanoscale, of amorphous carbon films and carbon–nitrogen films was that of Miyake and collaborators [135]. They measured the wear depth as function of the load of films deposited by reactive ion-plating. In the range of loads from 50 to 100 μN, they show that for nitrogen-containing carbon films the wear depth was on the order of one tenth of the depth measured in carbon films deposited using the same technique.

More recently, a study of sputtered carbon–nitrogen films was performed by Wiens and collaborators [136]. Argon–nitrogen atmospheres were used to sputter a graphite target and the substrate was not biased. For films with nitrogen content between 6 and 18 at.%, the scratch depth under loads of 5 μN decreases from 1 to 0.7 nm. The same test performed with loads of 3 μN shows a decrease of the scratch depth from 0.5 to 0.3 nm. The scratch test was done using a diamond tip and scanning velocity of 0.2 μm/s. Raman analysis carried out in these samples suggested that the films with higher scratch resistance have lower sp^2 content. The same group [137] performed a more detailed investigation and the wear behavior of a-CN films deposited by FCVA was studied. They found an opposite behavior for the scratch depth upon nitrogen incorporation into sputtered films and FCVA-deposited films. For sputtered films, nitrogen incorporation reduces the scratch depth and for FCVA-deposited films it results in an increase of the depth. Unfortunately the authors did not report the values either for elastic modulus of sputtered films or for the position of the G-band for nitrogen-incorporated FCVA-deposited films. The results are presented in Table 3.

Carbon–nitrogen films deposited by ion beam-assisted deposition were also studied [105]. The films have a nitrogen content of 11 at.% and the hardness is 22 GPa. The wear depth for 10 nm thick films submitted to a load of 10 μN is on the order of 1 nm. These values are higher than those obtained by Wiens et al. [125] for sputtered films with

Table 3. Nitrogen content, elastic modulus, Raman G-band position, and scratch depth for a-CN films deposited by sputtering and filtered cathodic vacuum arc techniques.

Deposition technique	Nitrogen content (at.%)	Elastic modulus (GPa)	G-band position (cm^{-1})	Scratch depth (Å)
FCVA	0	480	1558	0.7
FCVA	6	400	—	0.8
FCVA	12.5	150	—	1.5
FCVA	14.5	120	—	1.5
Sputtering	0	—	1584	4
Sputtering	6	—	1580	3.5
Sputtering	8.5	—	1579	3.7
Sputtering	14	—	1578	3.2
Sputtering	18	—	1575	3

Note: The wear test was performed using a diamond tip at a normal load of 2.8 μN [136, 137].

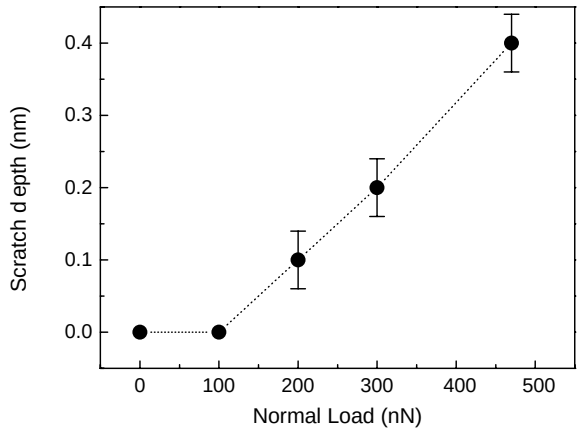


Figure 21. The scratch depth as a function of the normal load for FLC films [102]. The line is only to guide the eyes.

Table 4. Friction coefficient for several carbon based materials measured by FFM in air.

Materials	Friction coefficient
a-C films deposited by resistive evaporation [96]	0.45 ± 0.04
a-C films deposited by ion assisted sputtering [55][a]	0.28
a-C films deposited by ion assisted sputtering [114]	0.22 ± 0.02
a-C films deposited by laser ablation [102]	0.21 ± 0.03
Diamond [96]	0.25 ± 0.06
ta-C deposited by FCVA [112]	0.15 ± 0.03
ns-C (small clusters material) [114]	0.10 ± 0.04
ns-C (large clusters material) [114]	0.14 ± 0.05
FLC films [112]	0.12 ± 0.03
HOPG [102]	0.009 ± 0.003
HOPG [114]	0.004 ± 0.001
HOPG [96]	0.0012 ± 0.0009

[a] Experimental errors are not quoted by the authors.

Table 5. Friction coefficient for CN_x films (N content of the order of 10 at.%) measured by FFM in air.

Materials	Friction coefficient	
	a-C	a-CN
a-C:H films deposited by PECVD [106]	0.22 ± 0.05	0.25 ± 0.06
a-C films deposited by ion-assisted sputtering [55][a]	0.28	0.14
a-C films deposited by laser ablation [102]	0.21 ± 0.03	0.17 ± 0.03

[a] Experimental errors are not quoted by the authors.

similar nitrogen content. The possible explanation for this large dispersion of data can be found in the formation of the fullerene-like structure in nitrogen–carbon films deposited by sputtering and many other techniques [58]. In this case the high elasticity of the materials certainly influences the wear resistance of the films. It is possibly more important when low loads were applied.

We investigated the case of FLC films deposited using an unfiltered localized high-pressure arc discharge [112]. The scratch test was performed using a Si_3N_4 tip of $k = 0.77$ N/m and scanning velocity of 2 μm/s. The scratch depth as a function of the normal load is shown in Figure 21. The wear of the surface was verified only when we scratched the surface with normal forces of at least 200 nN. Increasing the normal force beyond this point leads to an increase of the wear depth. The resistance to scratching of FLC films may be attributed to the high elasticity nature of the films [3]. The tip may deform the surface while scanning, but this mechanical change is promptly restored due to the high film elasticity. Once a permanent surface modification is achieved the film wear is influenced by the energy dissipated at the tip–surface interface and increases with the applied normal load.

5. CONCLUSIONS

The development of friction force microscopy in the last years resulted in significant progress in the understanding of either the basic friction laws and in the tribological performance of several coatings. Among those coatings, hard carbon films occupy a special place.

Concerning the basic friction law, it was clear that in nanoscale the friction forces depended on the contact area and that the relation between friction and load forces was no longer linear. In addition, it was shown that the velocity dependence behavior of the friction forces could be understood as a competition between two mechanisms—the cohesive forces responsible for the stick and slip movement and the kinetics nucleation of water meniscus between the moving parts—resulting in a logarithmic dependence of friction with the scanning velocity.

The friction coefficients measured in air by friction force microscopy for several carbon-based materials are listed in Table 4. It was clear that, despite the large discrepancies among the different experimental reports, the friction coefficient for HOPG is at least 1 order of magnitude lower compared with the friction coefficient obtained from a-C films. Among those different films, the nanostructured carbon films and the fullerene-like carbon films have lower friction coefficients while diamond-like carbon films show

friction coefficients similar to the one measured in diamond surfaces.

The incorporation of hydrogen was also investigated, but it is very difficult to determine its role precisely. In fact, if on one side, the deposition of a-C:H films by PECVD in highly hydrogen diluted atmospheres resulted in a decrease of the friction coefficient due to the complete saturation of surface dangling bonds, the opposite effect was observed for a-C:H films deposited by sputtering, where the increase of the hydrogen content in the films resulted in higher friction coefficient. On the other side, the incorporation of nitrogen, as a general rule, induced a slight reduction in the friction coefficient as one can see in Table 5.

Concerning wear behavior, scratch tests performed on carbon films thicker than 10 nm revealed that the wear resistance of these materials increased with the film hardness. For carbon–nitrogen films, the situation is more complex and the dispersion of the experimental data can be attributed to high elastic recovery of the films. It is due to the presence of nitrogen in a hexagonal graphite lattice that favors the formation of pentagons, inducing curvature in the graphene planes, responsible for the elastic behavior of the films. It is of importance also for fullerene-like films. For these materials, when low loads are applied the elastic deformation of the films can drastically reduce the wear when the sample surface is scanned with an AFM tip.

GLOSSARY

Adhesion A consequence of the forces that act at the materials' contact area to causes their sticking.

Atomic force microscopy/microscope (AFM) A microscope able to observe surfaces at atomic and molecular scale by measuring the interaction forces between a sharp tip and a surface.

Friction Force opposing relative motion of two objects that are in contact.

Friction force microscopy/microscope (FFM) A microscope able to measure the friction forces between a sharp tip and a sample surface at atomic and molecular scale.

Fullerenes Molecular forms of pure carbon, with a cage-like structure of carbon atoms. The number of atoms ranges from 60 up to a few hundred carbon atoms.

Nanotribology The science and technology of interacting surfaces in relative motion at atomic and molecular scales. This includes friction, wear, adhesion, and lubrication.

Sputtering Method used to deposit thin films in which a plasma is used. The plasma ions are used to bombard the target atoms that will condense out on a substrate, placed inside the deposition chamber, depositing a film.

Thin films Films with thickness less than 100 μ.

Wear The removal of material from surfaces in relative motion as a result of surface contact or abrasive or chemical action.

ACKNOWLEDGMENTS

This work was partially supported by the Brazilian agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), and Fundação de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ).

REFERENCES

1. J. Robertson, *Mater. Sci. Eng., R* 37, 129 (2002).
2. F. L. Freire, Jr., *J. Non-Cryst. Solids* 304, 251 (2002).
3. G. A. J. Amaratunga, M. Chhowalla, C. J. Kiely, I. Alexandou, R. Aharonov, and R. M. Devenish, *Nature* 383, 321 (1996).
4. G. Binning, S. F. Quate, and C. Gerber, *Phys. Rev. Lett.* 56, 930 (1986).
5. R. Prioli, D. C. Reigada, and F. L. Freire, Jr., *J. Appl. Phys.* 87, 1118 (2000).
6. E. Meyer, R. M. Overney, K. Dransfeld, and T. Gyalog, "Nanoscience: Friction and Rheology on the Nanometer Scale." World Scientific, Singapore, 1996.
7. B. Bhushan, J. N. Israelachvill, and U. Landman, *Nature* 374, 607 (1995).
8. A. Grill, *Diamond Relat. Mater.* 10, 234 (2001).
9. S. Aisenberg and R. Chabot, *J. Appl. Phys.* 42, 2953 (1971).
10. Y. Lifshitz, *Diamond Relat. Mater.* 8, 1659 (1999).
11. Y. Lifshitz, S. R. Kasi, and J. W. Rabalais, *Phys. Rev. Lett.* 62, 1290 (1989).
12. C. A. Davies, *Thin Solid Films* 226, 30 (1993).
13. J. R. Robertson, *Diamond Relat. Mater.* 2, 984 (1993).
14. P. J. Fallon, V. S. Veerasamy, C. A. Davies, J. Robertson, G. A. J. Amaratunga, W. I. Milne, and J. Koskinen, *Phys. Rev. B* 48, 4777 (1993).
15. J. Robertson, *Prog. Solid State Chem.* 21, 199 (1991).
16. R. G. Lacerda and F. C. Marques, *Appl. Phys. Lett.* 73, 619 (1998).
17. S. S. Camargo, A. L. Baia Neto, R. A. Santos, F. L. Freire, Jr., R. Carius, and F. Finger, *Diamond Relat. Mater.* 7, 1155 (1998).
18. J. A. Theil, *J. Vac. Sci. Technol., B* 17, 2397 (1999).
19. K. Endo and T. Tatsumi, *J. Appl. Phys.* 78, 1370 (1995).
20. H. Yang, D. J. Tweet, Y. Ma, and T. Nguyen, *Appl. Phys. Lett.* 73, 1514 (1998).
21. H. Yokomichi, T. Hayashi, and A. Masuda, *Appl. Phys. Lett.* 72, 2704 (1998).
22. A. Grill, V. Patel, and C. Jahnes, *J. Electrochem. Soc.* 145, 1649 (1998).
23. T. E. Karis, G. W. Tyndall, D. Fenzel-Alexander, and M. S. Crowder, *J. Vac. Sci. Technol., A* 15, 2382 (1997).
24. C. Donnet, J. Fontaine, A. Grill, V. Patel, C. Jahnes, and M. Belin, *Surf. Coat. Technol.* 94–95, 531 (1997).
25. L. G. Jacobsohn, D. F. Franceschini, M. E. H. Maia da Costa, and F. L. Freire, Jr., *J. Vac. Sci. Technol., A* 18, 2230 (2000).
26. F. L. Freire, Jr., M. E. H. Maia da Costa, L. G. Jacobsohn, and D. F. Franceschini, *Diamond Relat. Mater.* 10, 125 (2001).
27. R. d'Agostino, R. Lamendola, P. Favia, and A. Giquel, *J. Vac. Sci. Technol., A* 12, 308 (1994).
28. H. Ji, A. Côté, D. Koshel, B. Terreault, G. Abel, P. Ducharme, G. Ross, S. Savoie, and M. Gagné, *Thin Solid Films* 405, 104 (2002).
29. F. L. Freire, Jr., *Jpn. J. Appl. Phys.* 36, 4886 (1997).
30. D. F. Franceschini, in "Thin Films and Nanostructures," Vol. 30, p. 217. Academic Press, New York, 2002.
31. A. Y. Liu and M. L. Cohen, *Science* 245, 841 (1989).
32. M. R. Wixon, *J. Am. Ceram. Soc.* 73, 1973 (1990).
33. J. H. Nguyen and R. Jeanloz, *Mater. Sci. Eng., A* 209, 23 (1996).
34. D. Marton, K. J. Boyd, and J. W. Rabalais, *Int. J. Mod. Phys., B* 9, 3527 (1995).
35. S. Muhl and J. M. Mendez, *Diamond Relat. Mater.* 8, 1809 (1999).
36. K. M. Yu, M. L. Cohen, E. E. Haller, W. L. Hansen, A. Y. Liu, and I. C. Wu, *Phys. Rev. B* 49, 5034 (1994).
37. C. Niu, Y. Z. Lu, and C. M. Lieber, *Science* 261, 334 (1993).
38. F. Wech, J. Widany, and Th. Frauenheim, *Phys. Rev. Lett.* 78, 3326 (1997).

39. D. F. Franceschini, F. L. Freire, Jr., and S. R. P. Silva, *Appl. Phys. Lett.* 68, 2645 (1996).
40. V. S. Veerasamy, J. Yuan, G. A. J. Amaratunga, W. I. Milne, K. W. R. Gilkes, M. Weiler, and L. M. Brown, *Phys. Rev. B* 48, 17954 (1993).
41. J. K. Walters, M. Kühn, C. Spaeth, H. Fischer, F. Richter, and R. J. Newport, *Phys. Rev. B* 56, 14315 (1997).
42. J. K. Walters, M. Kühn, C. Spaeth, H. Fischer, E. Dooryhee, and R. J. Newport, *J. Appl. Phys.* 83, 3529 (1998).
43. D. F. Franceschini, C. A. Achete, and F. L. Freire, Jr., *Appl. Phys. Lett.* 60, 3229 (1992).
44. S. Metin, J. H. Kaufman, D. D. Saperstein, J. C. Scotch, J. Heyman, and E. E. Haller, *J. Mater. Res.* 9, 396 (1994).
45. L. G. Jacobsohn, F. L. Freire, Jr., D. F. Franceschini, M. M. Lacerda, and G. Mariotto, *J. Vac. Sci. Technol., A* 17, 545 (1999).
46. S. R. P. Silva, J. Robertson, G. A. J. Amaratunga, B. Raferty, L. M. Brown, J. Schwan, D. F. Franceschini, and G. Mariotto, *J. Appl. Phys.* 81, 2626 (1997).
47. S. Rodil, N. A. Morrison, W. I. Milne, J. Robertson, V. Stolojan, and D. N. Jayawardane, *Diamond Relat. Mater.* 9, 524 (2000).
48. D. Li, X. Chu, S. C. Cheng, X. W. Lin, V. P. Dravid, Y. W. Chung, M. S. Wong, and W. D. Sproul, *Appl. Phys. Lett.* 67, 203 (1995).
49. D. Li, X. W. Lin, S. C. Cheng, V. P. Dravid, Y. W. Chung, M. S. Wong, and W. D. Sproul, *Appl. Phys. Lett.* 68, 1211 (1996).
50. M.-L. Wu, M. U. Guruz, W. A. Chiou, V. P. Dravid, Y. W. Chung, S. Anders, F. L. Freire, Jr., and G. Mariotto, *Appl. Phys. Lett.* 76, 2692 (2000).
51. D. Li, Y. W. Chung, M. S. Wong, and W. D. Sproul, *J. Appl. Phys.* 74, 219 (1993).
52. W. C. Chan, B. Zhou, Y. W. Chung, C. S. Lee, and S. T. Lee, *J. Vac. Sci. Technol., A* 16, 1907 (1998).
53. B. Wei, B. Zhang, and K. E. Johnson, *J. Appl. Phys.* 83, 2491 (1998).
54. T. A. Yeh, C. L. Lin, J. M. Sivertsen, and J. H. Judy, *J. Magn. Magn. Mater.* 120, 314 (1993).
55. A. Khrushudov, K. Kato, and S. Daisuke, *J. Vac. Sci. Technol., A* 14, 2935 (1996).
56. M. U. Guruz, V. P. Dravid, Y. W. Chung, M. M. Lacerda, C. S. Bhatia, Y. H. Yu, and S. C. Lee, *Thin Solid Films* 381, 6 (2001).
57. J. J. Cuomo, P. A. Leary, D. Yu, W. Reuter, and M. Frisch, *J. Vac. Sci. Technol.* 16, 299 (1979).
58. H. Sjöström, S. Stafström, M. Boman, and J. E. Sundgren, *Phys. Rev. Lett.* 75, 1336 (1995).
59. M. P. Siegal, D. R. Tallant, P. N. Provencio, D. L. Overmyer, and R. L. Simpson, *Appl. Phys. Lett.* 76, 3052 (2000).
60. P. R. Goglia, J. Berkowitz, J. Hoehn, A. Xidis, and L. Stover, *Diamond Relat. Mater.* 10, 271 (2001).
61. A. Khurshudov and R. J. Waltman, *Wear* 251, 1124 (2001).
62. D. J. Li, M. U. Guruz, C. S. Bhatia, and Y. W. Chung, *Appl. Phys. Lett.* 81, 1113 (2002).
63. J. R. Robertson, *Thin Solid Films* 383, 81 (2001).
64. E. Liu, B. Blanpain, and J. P. Celis, *Wear* 192, 141 (1996).
65. R. G. Cain, M. G. Reitsma, S. Biggs, and N. W. Page, *Rev. Sci. Instrum.* 72, 3304 (2001).
66. D. Dowson, "History of Tribology." Longman, London, 1979.
67. C. M. Mate, G. M. McClelland, R. Erlandson, and S. Chiang, *Phys. Rev. Lett.* 59, 1942 (1987).
68. G. A. Tomlinson, *Philos. Mag. Ser.* 7, 905 (1929).
69. G. V. Dedkov, *Phys. Status Solidi A* 179, 3 (2000).
70. H. Hölscher, U. D. Schartz, O. Zwörner, and R. Wiesendanger, *Phys. Rev. B* 57, 2477 (1998).
71. N. Sasaki, K. Kobayashi, and M. Tsukada, *Phys. Rev. B* 54, 2138 (1996).
72. N. Sasaki, M. Tsukada, S. Fugisawa, Y. Sugawara, S. Morita, and K. Kobayashi, *Phys. Rev. B* 57, 3785 (1998).
73. N. Sasaki, K. Kobayashi, and M. Tsukada, *Surf. Sci.* 92, 357 (1996).
74. A. M. F. Rivas, R. M. Zamora, and R. Prioli, Ultramicroscopy, in press.
75. M. A. Lantz, S. J. O'Shea, M. E. Welland, and K. L. Johnson, *Phys. Rev. B* 55, 10776 (1997).
76. C. A. J. Putmann, M. Igarshi, and R. Kaneko, *Appl. Phys. Lett.* 66, 3221 (1995).
77. A. Opitz, S. I.-U. Ahmed, J. A. Schaefer, and M. Scherge, *Surf. Sci.* 504, 199 (2002).
78. E. Riedo, F. Levy, and H. Brune, *Phys. Rev. Lett.* 88, 1855 (2002).
79. J. S. Helman, W. Baltensperger, and J. A. Holyst, *Phys. Rev. B* 49, 3841 (1994).
80. F. Family, H. G. E. Hentschel, and Y. Braiman, *J. Phys. Chem. B* 104, 3984 (2000).
81. R. Prioli, A. M. F. Rivas, F. L. Freire, Jr., and A. O. Caride, *Appl. Phys. A,* in press.
82. T. Bouhacina, J. P. Aime, S. Gauthier, D. Michel, and V. Heroguez, *Phys. Rev. B* 56, 7694 (1997).
83. E. Gnecco, R. Bennewitz, T. Gyalog, Ch. Loppacher, M. Bammerlin, E. Meyer, and H.-J. Güntherodt, *Phys. Rev. Lett.* 84, 1172 (2000).
84. A. J. Lichtenberg and M. A. Liebermann, "Regular and Stochastic Motion." Springer-Verlag, New York, 1981.
85. K. L. Johnson, K. Kendall, and A. D. Roberts, *Proc. R. Soc. London, Ser. A* 324, 301 (1971).
86. B. V. Derjaguin, V. M. Muller, and Y. P. Toporov, *J. Colloid Interface Sci.* 53, 314 (1975).
87. H. Hertz, *J. Reine Angew. Math.* 92, 156 (1881).
88. D. Maugis, *J. Colloid Interface Sci.* 53, 243 (1992).
89. D. Tabor, *J. Colloid Interface Sci.* 58, 2 (1977).
90. L. O. Heim, J. B. M. Preuss, and H. J. Butt, *Phys. Rev. Lett.* 83, 3328 (1999).
91. A. N. Israelachvili, "Intermolecular and Surface Forces with Application to Colloidal and Biological Systems", p 10. Academic Press, London, 1985.
92. B. Bushan and C. Dandavate, *J. Appl. Phys.* 87, 1201 (2000).
93. A. Greenwood and B. J. P. Williamson, *Proc. R. Soc. London, Ser. A* 295, 300 (1966).
94. A. Greenwood and J. H. Tripp, *Trans ASME Ser. E, J. Appl. Mech.* 34, 153 (1967).
95. A. Greenwood, *Trans. ASME Ser. F, J. Lubrication Technol.* 89, 81 (1967).
96. U. D. Schwarz, O. Zwöner, P. Köster, and R. Wiesendanger, *Phys. Rev. B* 56, 6987 (1997).
97. M. Binggeli and C. M. Mate, *Appl. Phys. Lett.* 65, 415 (1994).
98. S. S. Perry, C. Mathew Mate, R. L. White, and G. A. Samorjai, *IEEE Trans. Mag.* 32, 115 (1996).
99. A. Grill, *Surf. Coat. Technol.* 94/95, 507 (1997).
100. M. Cieplak, E. D. Smith, and M. O. Robbins, *Science* 265, 1209 (1994).
101. B. N. J. Persson, *Phys. Rev. B* 44, 3277 (1991).
102. E. Riedo, J. Chevrier, F. Comin, and H. Brune, *Surf. Sci.* 477, 25 (2001).
103. B. Bhushan and S. Sundararajan, *Acta Mater.* 46, 3793 (1998).
104. S. Sundararajan and B. Bushan, *J. Appl. Phys.* 88, 4825 (2000).
105. M. Bai, K. Kato, N. Umehara, Y. Miyake, J. Xu, and H. Tokisue, *Surf. Coat. Technol.* 126, 181 (2000).
106. R. Prioli, S. I. Zanette, A. O. Caride, F. L. Freire, Jr., and D. F. Franceschini, *J. Vac. Sci. Technol., A* 14, 2351 (1996).
107. D. F. Franceschini, F. L. Freire, Jr., and S. R. P. Silva, *Appl. Phys. Lett.* 68, 2645 (1996).
108. Y. Lifshitz, G. D. Lempert, and E. Grossman, *Phys. Rev. Lett.* 72, 2753 (1994).
109. S. Sundararajan and B. Bhushan, *Wear* 225–229, 678 (1999).
110. O. Zwörner, H. Hölscher, U. D. Schwarz, and R. Wiesendanger, *Appl. Phys. A* 66, S263 (1998).
111. K. I. Schiffmann, *Tribol. Lett.* 5, 109 (1998).

112. F. L. Freire, Jr. and R. Prioli, Technical Report PUC-VDG 01, 2001.
113. E. Liu, B. Blanpain, J. P. Celis, and J. R. Roos, *J. Appl. Phys.* 84, 4859 (1998).
114. R. Buzio, E. Gnecco, G. Boragno, and U. Valbusa, *Carbon* 40, 883 (2002).
115. D. Donadio, L. Colombo, P. Milani, and G. Benedek, *Phys. Rev. Lett.* 83, 776 (1999).
116. H. Liu, S. Imad-Uddin Ahmed, and M. Scherge, *Thin Solid Films* 381, 135 (2001).
117. F. L. Freire, Jr., L. G. Jacobsohn, M. E. H. Maia da Costa, and R. Prioli, Technical Report PUC-VDG 01, 2002.
118. M. Binggeli and C. M. Mate, *J. Vac. Sci. Technol., B* 13, 1312 (1995).
119. A. Erdemir, *Surf. Coat. Technol.* 146–147, 292 (2001).
120. J. Fontaine, C. Donnet, A. Grill, and T. LeMogne, *Surf. Coat. Technol.* 146–147, 286 (2001).
121. T.-H. Fang, C.-I. Weng, J.-G. Chang, and C.-C. Hwang, *Thin Solid Films* 396, 166 (2001).
122. A. Erdemir, I. B. Nilufer, O. L. Erylmaz, M. Bescheliesser, and G. R. Fenske, *Surf. Coat. Technol.* 120/121, 589 (1999).
123. Z. Jiang, C. J. Lu, D. B. Bogy, C. S. Bathia, and T. Miyamoto, *Thin Solid Films* 258, 75 (1995).
124. B. Bhushan, *Wear* 251, 1105 (2001).
125. A. Wiens, G. Persch-Schuy, M. Volgelgasang, and U. Hartmann, *Appl. Phys. Lett.* 75, 1077 (1999).
126. E. V. Anoikin, M. M. Yang, J. L. Chao, J. R. Elings, and D. W. Brown, *J. Vac. Sci. Technol., A* 16, 1741 (1998).
127. G. Mariotto, C. A. Achete, and F. L. Freire, Jr., *Thin Solid Films* 241, 255 (1994).
128. T.-H. Fang, C.-I. Weng, and M.-J. Chiang, *Diamond Relat. Mater.* 11, 1653 (2002).
129. S. Sundararajan and B. Bhushan, *Wear* 217, 251 (1998).
130. B. K. Gupta and B. Bushan, *Thin Solid Films* 270, 391 (1995).
131. E. Martinez, J. L. Andújar, M. C. Polo, J. Esteve, J. Robertson, and W. I. Milne, *Diamond Relat. Mater.* 10, 145 (2001).
132. B. Bushan and V. N. Koinkar, *J. Appl. Phys.* 75, 5741 (1994).
133. D. Sheeja, B. K. Tay, S. P. Lau, X. Shi, and X. Ding, *Surf. Coat. Technol.* 132, 228 (2000).
134. C. A. Davis, G. A. J. Amartunga, and K. M. Knowles, *Phys. Rev. Lett.* 80, 3280 (1998).
135. S. Miyake, S. Watanabe, H. Miyazawa, M. Murakawa, R. Kaneko, and T. Miyamoto, *Appl. Phys. Lett.* 65, 3206 (1994).
136. A. Wiens, G. Persch-Schuy, R. Hartmann, and P. Loeris, *J. Vac. Sci. Technol., A* 18, 2023 (2000).
137. A. Wiens, M. Neuhäuser, H.-H. Schneider, G. Persch-Schuy, J. Windeln, T. Witke, and U. Hartmann, *Diamond Relat. Mater.* 10, 1024 (2001).

ENN Encyclopedia of Nanoscience and Nanotechnology
www.aspbs.com/enn

Nanotubes for Nanoelectronics

Zhi Chen

University of Kentucky, Lexington, Kentucky, USA

CONTENTS

1. INTRODUCTION

Carbon nanotubes (CNTs) have emerged as a viable electronic material for molecular electronic devices because of their unique physical and electrical properties [1–7]. For example, nanotubes have a lightweight and record-high elastic modulus, and they are predicted to be by far the strongest fibers that can be made. Their high strength and high flexibility are unique mechanical properties. They also have amazing electrical properties. The electronic properties depend drastically on both the diameter and the chirality of the hexagonal carbon lattice along the tube [8–10].

Carbon nanotubes were discovered by Iijima in 1991 at the NEC Fundamental Research Laboratory in Tsukuba, Japan [1]. Using a transmission electron microscope (TEM), he found carbon tubes consisting of multiple shells (see Fig. 1). These early carbon tubes are called multiwall nanotubes (MWNTs). Since then, extensive research has been focused on synthesis and characterization of carbon nanotubes. In 1993, Ijima's group [11] and Bethune et al. [12] at IBM Almaden Research Center at San Jose, California, synthesized carbon nanotubes with a single shell, called single-wall nanotubes (SWNTs). Because of their simple and well-defined structure [13], the single-wall nanotubes serve as model systems for theoretical calculations and for critical experimental studies. Since then, the physical and electrical properties of carbon nanotubes have been studied extensively. Only a few years ago, people began to utilize carbon nanotubes' unique electrical properties for electron device applications. Up to now, there has been no extensive review to cover the progress in nanoelectronic devices using carbon nanotubesy.

In this chapter, I aim to present extensive review on progress in electronic structure and transport properties of carbon nanotubes and nanoelectronic devices based on carbon nanotubes ranging from quantum transport to field-effect transistors. This chapter is organized as follows. It begins with brief review of the progress in micro- and nanoelectronic devices prior to carbon nanotubes (Section 2). Then the synthesis and physical properties of carbon nanotubes will be discussed in Section 3. Nanoelectronic devices based on carbon nanotubes including single-electron transistors (SETs), field-effect transistors, logic gates, and memory devices will be reviewed in Sections 4 and 5. The chemical doping, junctions, and metal–nanotube contacts will be described in Section 6. Finally, nanofabrication based on carbon nanotubes including controlled growth and selective placement of nanotubes on patterned Si substrates will be reviewed in Section 7. Since the discovery of carbon nanotubes, over 1000 papers on carbon nanotubes have been published. It is unlikely that every paper will be included in this chapter, because most papers dealt with synthesis, physical, and chemical properties of nanotubes. In this chapter, I will focus on electrical properties of nanotubes, nanoelectronic devices constructed with nanotubes, and nanotube-based nanofabrication.

2. PROGRESS IN NANOELECTRONICS PRIOR TO NANOTUBES

In 1959 Richard Feynman delivered his famous lecture, "There is Plenty of Room at the Bottom." He presented a vision of exciting new discoveries if one could fabricate materials and devices at the atomic and molecular scale. It was not until 1980s that instruments such as scanning tunneling microscopes (STM), atomic force microscopes (AFM), and near-field microscopes were invented. These instruments

ISBN: 1-58883-063-2/$35.00

Encyclopedia of Nanoscience and Nanotechnology
Edited by H. S. Nalwa
Volume 7: Pages (919–942)

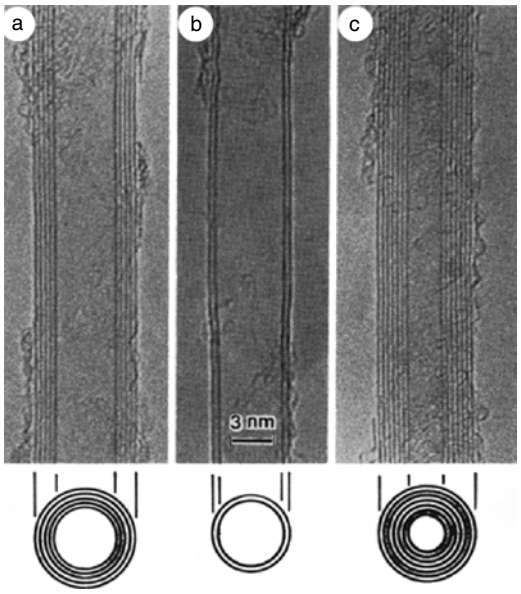


Figure 1. Cross-section images of carbon nanotubes by a high-resolution TEM. Reprinted with permission from [1], S. Iijima, *Nature* 354, 56 (1991). © 1991, Macmillan Magazines Ltd.

provide "eyes" and "fingers" required for nanostructure measurement and manipulation [14]. The driving force for nanoelectronics is the scaling of microelectronic devices to nanoscale, which is the engine for modern information revolution. The microelectronics revolution began in 1947 when John Bardeen, Walter H. Brattain, and William Shockley of Bell Telephone Laboratories invented the first solid state transistor, the Ge point-contact transistor [15]. Solid state transistors had far superior performance, much lower power consumption, and much smaller size than vacuum triodes. People began to produce individual solid state components to replace the vacuum tubes in circuits in a few years after the invention. In 1958, Jack Kilby of Texas Instruments conceived a concept for fabrication of the entire circuit including components and interconnect wires on a single silicon substrate [16]. In 1959, Robert Noyce of Fairchild Semiconductor individually conceived a similar idea [15]. This concept has evolved into today's very-large-scale integrated circuits or "microchips," which consist of millions of transistors and interconnect wires on a single silicon substrate. The major driving force for this revolutionary progress is miniaturization of transistors and wires from tens of micrometers in the 1960s to today's tens of nanometers. The scaling down of transistor and interconnect wires led to more and more transistors being incorporated into a single silicon chip, resulting in faster and more powerful computer chips. With the success in microelctronics and of the semiconductor industry, it is natural to consider extending the micrometer size devices to nanometer size devices.

2.1. STM/AFM-Based Nanofabrication

There was not much progress in nanoelectronics until the 1980s. In 1981, the scanning tunneling microscope, invented by G. K. Binnig and H. Rohrer of IBM Zurich Research Laboratory, produced the first images of individual atoms on the surfaces of materials [17]. Later the STM was used to modify surfaces at the nanometer scale and the manipulation and positioning of single atoms on surface was achieved [18]. In 1993, Crommie et al. [19] eloquently demonstrated with a "quantum corral" where the STM can be used not only to characterize the electronic structure of materials on a truly quantum scale, but also to modify this quantum structure. This suggested that the STM and AFM might be used for atom-by-atom control of materials modification, leading to atomic resolution. This potential motivated researchers to attempt to use STM and AFM for nanolithography [20–25]. This task has not been easy due to the irreproducibility of the modifications, the slow "write" speed, and the difficulty of transferring such fine manipulations into functioning semiconductor devices [21]. Until now, reliable fabrication and robust pattern transfer for linewidths below 10 nm has not been achieved yet.

2.2. Metal-Oxide-Semiconductor Field-Effect Transistors

As early as the 1920s and 1930s, a concept for amplifying devices based on the so-called *field effect* was proposed with little understanding of the physical phenomena [26, 27]. After 30 years, the field-effect transistor based on the SiO_2/Si structure finally became practical [28]. Since that time, the metal-oxide-semiconductor field-effect transistor (MOSFET) has been incorporated into integrated circuits and has grown to be the most important device in the semiconductor industry ranging from memory chips, microprocessors, and many other communication chips. Figure 2a shows the structure of an *n*-channel MOSFET. It consists of *p*-type Si, heavily doped n^+ source and drain, and an insulated gate. When the gate voltage is zero, the region underneath the gate oxide is *p*-type. There are two pn^+ junctions near the source and the drain. When applying a voltage across the source and the drain, either of the two pn^+ junctions is reverse biased. Thus the transistor conducts no current. When the gate voltage is larger than zero but less than its threshold voltage, depletion happens underneath the gate oxide, which is still not conductive. When the gate voltage is larger than its threshold voltage, an inversion layer (*n*-type) is induced underneath the gate oxide, which forms a conduction channel connecting both source and drain. Thus the transistor conducts current. The typical *I–V* curves for an *n*-MOSFET with a gate length of 0.25 μm and a width of 15 μm are shown in Figure 2b. The driving force behind this

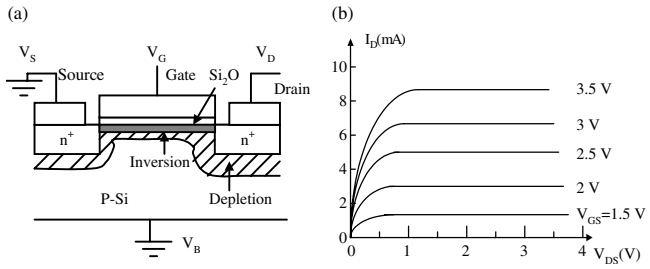


Figure 2. (a) Schematic structure of a Si MOSFET used in various microchips for digital signal processing and (b) its current–voltage characteristics.

remarkable development is the cost reduction and performance enhancement of integrated circuits (ICs) due to the continuous miniaturization of transistors and interconnect wires. The continuous scaling of transistors and the increase in wafer size has been and will continue to be the trend for the semiconductor industry. The transistor gate length (feature size) has been dramatically reduced for the past three decades [29, 30]. The lateral feature size or linewidth of a transistor has been shrunk by almost 50 times from about 10 μm to the 100 nm range, allowing over 10,000 times more transistors to be integrated on a single chip. Microprocessors have evolved from their 0.1 MHz ancestors used for watches and calculators in the early 1970s to the current 2 GHz engines for personal computers. Memories have grown from 1-Kb pioneers used almost exclusively for mass storage in central computers to the 128/256-Mb dynamic random access memory commonly used in personal computers today. Most advanced chips on the market now have feature sizes of 100 nm. According to the Semiconductor Industry Association's *International Technology Roadmap for Semiconductors* [31], the feature sizes for lithography were projected as follows: *130 nm in 2001, 100 nm in 2003, 80 nm in 2005, 35 nm in 2007, 45 nm in 2010, 32 nm in 2013, and 22 nm in 2016*. The wafer size has increased from 2 inches in the early 1970s to the current 12 inches. Therefore, the performance of ICs has been dramatically improved, and the cost for manufacturing has been dramatically reduced. However, the device scaling is approaching its limit. It was suggested that MOS device scaling might not be extended to below 10 nm because of physical limits such as power dissipation caused by leakage current through tunneling [32–34]. In addition, once electronic devices approach the nano- and molecular scale, the bulk properties of solids are replaced by the quantum-mechanic properties of a relatively few atoms such as energy quantization and tunneling. It is therefore important to search for alternative devices for Si MOS devices.

2.3. Quantum-Effect Devices

In order for a transistor-like device to operate on the nano- and molecular scale, it would be advantageous if it operated based on quantum mechanical effects. Even though the study of single-electron charging effects with granular metallic systems dates back to the 1950s [35, 36], it was the research of Likharev in 1988 that laid much of the groundwork for understanding single-charge transport in nanoscale tunnel junctions [37, 38]. When a small conductor (island) is initially neutral, it does not generate any appreciable electric field beyond its border. In this state, a weak external force (due to electric field) may bring in an additional electron from outside. The net charge in the island is $(-e)$ and the resulting electric field repulses the following electrons which might be added. In order to have an electron to be added it needs to overcome the charging energy and its kinetic energy. Thus, the electron additional energy E_a is given by

$$E_a = E_c + E_k$$

Here E_c is the charging energy that is given by

$$E_c = \frac{e^2}{C}$$

where e is the charge of electron and C is the capacitance of the conductive island. The kinetic energy is expressed as

$$E_k = \frac{1}{g(\varepsilon_F)V}$$

where $g(\varepsilon_F)$ is the density of states on the Fermi surface and V is the volume of the island.

For the island diameter > 2 nm, $E_c \gg E_k$. Thus

$$E_a \approx E_c = \frac{e^2}{2C}$$

For the 100 nm-scale island, E_a is of the order of 1 meV, corresponding to $\sim$10 K in temperature. If the island size is reduced to below 10 nm, E_a approaches 100 meV, and some single-electron effects become visible at room temperature.

Among quantum-effect devices, single-electron transistors have been most extensively studied [38–40]. The basic structure of a SET is shown in Figure 3a. An island or quantum dot is placed between two electrodes (source and drain) with the third electrode (gate) is placed by its side. When a voltage is applied, an electron tunnels onto the island and the charging energy is increased by $E_a \approx e^2/2C$ and this increase acts as a barrier to the transfer of any further electrons. At small source–drain voltage, there is no current. The I–V characteristic is shown in Figure 3b. The current is blocked from $-V_c$ to V_c, called *Coulomb blockade*. When the source–drain voltage is increased and reaches a level greater than V_c, where the energy barrier is eliminated, electrons can cross the island and the current increases with the applied voltage. The threshold voltage V_c is a periodical function of gate voltage.

Coulomb charging effects were originally observed in metallic film by Gorter in 1951 [35]. The first successful metallic single-electron transistor was made by Fulton and Dolan in 1987 [36]. They used a relatively simple technique in which two layers of aluminum were evaporated *in-situ* from two angles through the same suspended mask formed by direct e-beam writing. Since then, single-electron transistors have been demonstrated in numerous experiments using a wide variety of device geometry, materials, and techniques. SETs based on metallic nanodots were fabricated. Chen et al. [41] reported that SETs with metal dots of 20–30 nm and gaps of 20–30 nm were fabricated by ionized beam evaporation. The electrical results showed clear Coulomb blockade at temperature as high as 77 K. Novel lateral metallic SETs can be based on gold colloidal particles. These particles are very uniform in size and can be obtained in a range

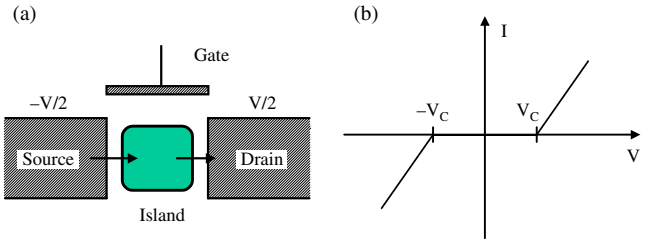


Figure 3. (a) Schematic structure of a capacitively coupled single-electron transistors and (b) its source–drain dc I–V curves.

of sizes. The devices can be fabricated by placing the particles in the gap between the source and drain. Klein et al. [42] obtained the Coulomb blockade characteristics based on gold colloidal particles at 77 K. Nakamura et al. [43] fabricated Al-based SETs, which operated at 100 K. Shirakashi et al. [44] fabricated Nb/Nb oxide-based SETs at room temperature ($T = 298$ K). The further reduction of the tunnel junction is performed by scanning probe microscope (SPM)-based anodic oxidation. The Coulomb blockade characteristics were clearly shown at room temperature.

Single-electron effects have also been observed in a number of semiconductor based structures. The first device was realized by squeezing the two-dimensional electron gas (2DEG) formed at the AlGaAs/GaAs heterostructure [45]. It consists of a pattern of metal gates evaporated onto the semiconductor surface. The voltages were applied to the gates to squeeze the 2DEG so that islands and tunnel barriers were formed. Coulomb blockade can be observed if the regions are sufficiently squeezed. Coulomb blockade effects have also been demonstrated on silicon based devices. A number of experiments have been reported on structures based on silicon-on-insulator (SOI). This is particularly important because silicon processing technology is the mainstream technology in the semiconductor industry. The silicon based SET technology can be easily integrated into the mainstream technology once it is successful. Ali and Ahmed [46] showed the first SOI-based SETs with clear Coulomb blockade. Leobandung et al. [47] also demonstrated silicon quantum-dot transistors with a 40-nm dot. The Coulomb blockade was clearly seen at temperature of 100 K. Takahashi et al. [48] and Kurihara et al. [49] scaled the silicon-based SETs to make significantly smaller islands and obtained Coulomb oscillation at temperatures approaching room temperature. Zhuang et al. [50] reported fabrication of silicon quantum-dot transistors with a dot of $\sim$12 nm with a clear Coulomb blockade at 300 K. Guo et al. [51] successfully fabricated a silicon single-electron transistor memory, which operated at room temperature. The memory is a floating gate MOS transistor in silicon with a channel width ($\sim$10 nm) smaller than the Debye screening length of a single electron and a nanoscale polysilicon dot (7 nm $\times$ 7 nm) as the floating gate embedded between the channel and the control gate. Storing one electron on the floating gate screens the entire channel from the potential on the control gate and leads to a discrete shift in the threshold voltage, a staircase relation between the charging voltage and the shift. SETs based on Si nanowires have also been demonstrated [52]. It was shown that quantum wires with a large length to width ratio show clear Coulomb oscillations at temperatures up to 77 K.

3. SYNTHESIS AND ELECTRICAL PROPERTIES OF CARBON NANOTUBES

3.1. Synthesis of Carbon Nanotubes

Although this chapter focuses on nanoelectronic devices, I still cover some of synthesis methods and approaches which may be helpful for interested readers. It should be pointed out that I am unable to include all the papers in synthesis because extensive research has been conducted for growth of carbon nanotubes.

Although multiwall carbon nanotubes were discovered in 1991 by Iijima, it is quite likely that such MWNTs were produced as early as the 1970s during research on carbon fibers [2]. The multiwall carbon nanotubes discovered in 1991 were obtained from the fullerene soot produced in an arc discharge [1]. As early as in 1986, Saito of the University of Kentucky studied the soot produced by a candle-like methane flame. A quartz fiber was inserted into the flame from the side and left there for a certain period of time [53]. When the fiber was raised to a certain height, a smooth film was coated on the fiber surface, which appeared to be brown in color. With increase of the sampling height beyond the critical height, the color of the deposited material changed from brown to black, and its surface appearance also changed from smooth to a rough and bumpy structure. The SEM analysis identified the rough surface material to be soot and the initial light brown material from the methane flame to be polyhedral-shaped crystal-like particles [54]. The deposit from the acetylene flame had a spider-web shape of entangled, long, narrow diameter strings, of which the detailed structure remained unknown until recent TEM study [55]. It was a surprise that the recent TEM study showed that the entangled long strings synthesized in 1986 are *carbon nanotubes*, which were discovered later by Iijima in 1991. Recently Saito's group repeated his 1986 experiments and synthesized CNTs using methane flames [56] and ethylene flames [57]. In 1993, Iijima and Ichihashi [11] synthesized single-wall carbon nanotubes of 1-nm diameter. Bethune and co-workers [12] also, at the same time, synthesized single-wall carbon nanotubes using cobalt catalyst. The research on carbon nanotubes really took off when Smalley and co-workers at Rice University found a laser ablation technique that could produce single-wall carbon nanotubes at yields up to 80% instead of the few percent yields of early experiments [58–60]. Kong and co-workers [61] at Stanford University used a chemical vapor deposition (CVD) technique to grow carbon nanotubes by decomposing an organic gas over a substrate covered with metal catalyst particles. The CVD approach has the potential for making possible large-scale production of nanotubes and growth of nanotubes at specific sites on patterned Si substrates [62, 63].

3.2. Electronic Structures of Carbon Nanotubes

Just in one year after the discovery of carbon nanotubes, their electronic structures were theoretically studied based on local-density-functional calculation [8], tight-binding band-structure calculation [9, 10, 64]. Figure 4a shows how to construct a carbon nanotube by wrapping up a single sheet of graphite such that two equivalent sites of the hexagonal lattice coincide; that is, point C coincides with the origin (0, 0) [65]. The wrapping vector $\mathbf{C}$, which defines the relative location of the two sites, is specified by a pair of integers (n, m) that relate $\mathbf{C}$ to the two unit vectors $\mathbf{a}_1$ and $\mathbf{a}_2$ ($\mathbf{C} = n\mathbf{a}_1 + m\mathbf{a}_2$). A tube is called "armchair" if n equals m, and "zigzag" in the case $m = 0$. All other tubes are of the "chiral" type with a finite wrapping angle

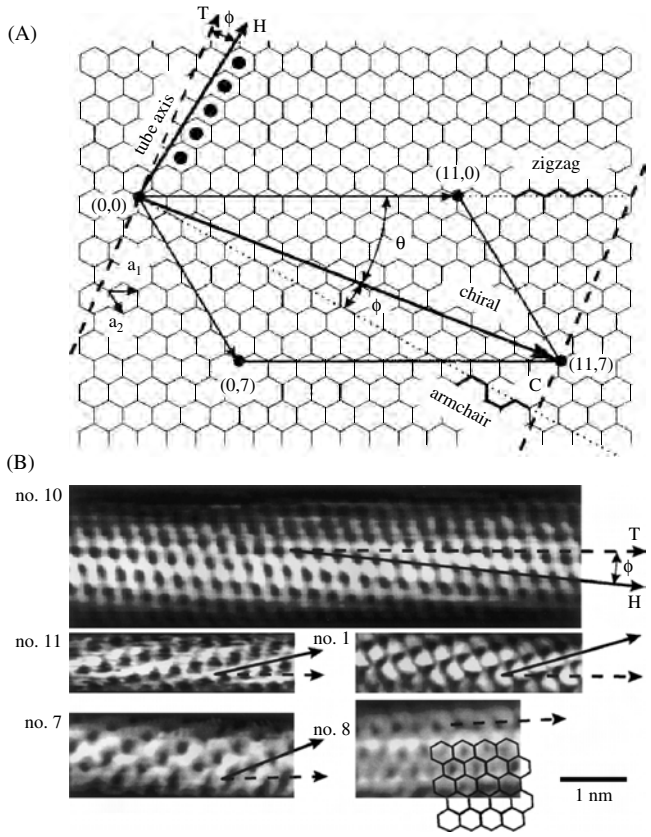


Figure 4. Relation between the hexagonal carbon lattice and the chirality of carbon nanotubes. (A) Construction of a carbon nanotube from a single graphene sheet by rolling up the sheet along the wrapping vector C. (B) Atomically resolved STM images of individually single-wall carbon nanotubes showing chirality. Reprinted with permission from [65], J. W. G. Wildoer et al., *Nature* 391, 59 (1998). © 1998, Macmillan Magazines Ltd.

ϕ ($0° < \phi < 30°$). Figure 4b shows the STM images of single-wall carbon nanotubes [65]. Tube 10 has a chiral angle $\phi = 7°$ and a diameter $d = 1.3$ nm, which corresponds to the (11, 7) type of panel A. The dependence of the electronic structure of nanotubes on the tube indices (n, m) can be understood by taking the two-dimensional graphene sheet as a starting point. In the circumferential direction (along **C**), the periodic boundary conditions $\mathbf{C} \cdot \mathbf{k} = 2\pi q$ can be applied, where **k** is the wave vector and q is an integer. This leads to a set of allowed values for k, which can be substituted into the dispersion relations for the tube, with q representing the various modes. Electronic energy band structure calculations [3, 8–10, 64] predicted that armchair ($n = m$) tubes behave like metallic. For all other tubes (chiral and zigzag) there exist two possibilities. If $(n - m)/3$ is an integer, tubes are expected to be metallic, and if $(n - m)/3$ is not an integer, tubes are predicted to be semiconducting with an energy gap depending on the diameter. The energy gap can be expressed as $E_{\text{gap}} = 2\gamma_0 a_{\text{C–C}}/d$, where γ_0 is the C–C tight-binding overlap energy, $a_{\text{C–C}}$ is the nearest neighbor C–C distance (0.142 nm), and d is the diameter. Figure 5 shows the calculated energy band structure of zigzag nanotubes (12, 0) in (c) and (13, 0) in (d). [The geometric structure of the tubes and the first Brillouin zone of a graphene sheet are

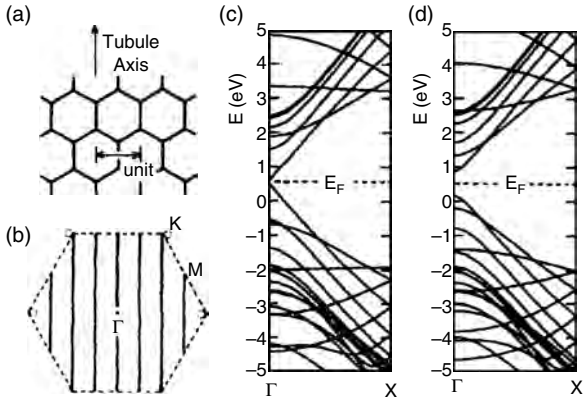


Figure 5. (a) The geometric configuration for a single-wall carbon nanotube $(n, 0)$. (b) The first Brilluoin zone of a graphite sheet and the wave vector allowed by the periodic boundary condition along the circumference for $n = 6$ (solid lines). Band structures of (c) (12,0) and (d) (13,0) single-wall nanotubes. Reprinted with permission from [9], N. Hamada et al., *Phys. Rev. Lett.* 68, 1579 (1992). © 1992, American Physical Society.

shown in (a) and (b).] The tube (12,0) is metallic, satisfying the condition of $(n - m)/3$ being integer; and the tube (13,0) is semiconducting with an energy gap of 0.697 eV, which falls into the category of $(n - m)/3$ being noninteger. Figure 6 shows the calculated energy band gaps of tubes $(n, 0)$ with $n = 6$–15. For $n = 6$, 9, 12, 15 [i.e., $(n - m)/3$ is integer], energy gaps are almost zero, and for $n = 7, 8, 10, 11, 13, 14$, [i.e. $(n - m)/3$ is noninteger], the energy gaps are ranging from 0.6 to 1.2 eV. Figure 7 shows the calculated one-dimensional (1D) electronic density of states for (a) a (9,0) nanotube and (b) a (10,0) nanotube [66]. The 1D density of state (DOS) of both nanotubes shows a series of spikes. Each spike corresponds to the energy threshold for an electronic subband caused by the quantum confinement of electrons in the radial and circumferential directions of nanotubes. The (9,0) nanotube is metallic and the (10,0) tube is semiconducting.

Experimental measurements [65, 67] of the energy bands of nanotubes confirmed these theoretical calculations. Figure 8a shows a selection of I–V curves obtained by scanning tunneling spectroscopy (STS) on different tubes [65]. Most curves show a low conductance at low bias, followed by several kinks at larger bias voltages. Figure 8b shows

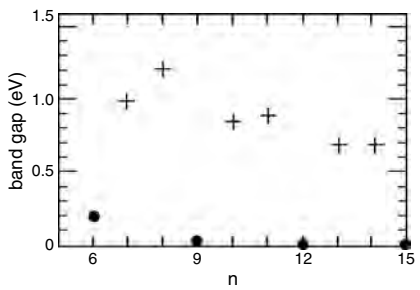


Figure 6. The energy bandgap as a function of the number of hexagons on the circumference for a single-wall nanotube $(n, 0)$. Reprinted with permission from [9], N. Hamada et al., *Phys. Rev. Lett.* 68, 1579 (1992). © 1992, American Physical Society.

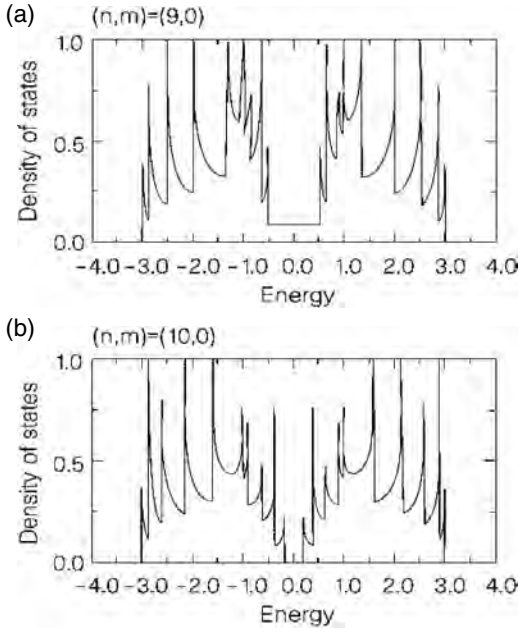


Figure 7. Calculated one-dimensional electronic density of states for (a) a (9,0) nanotube and (b) a (10,0) nanotube. Reprinted with permission from [66], M. S. Dresshaus, *Nature* 391, 19 (1998). © 1998, Macmillan Magazines Ltd.

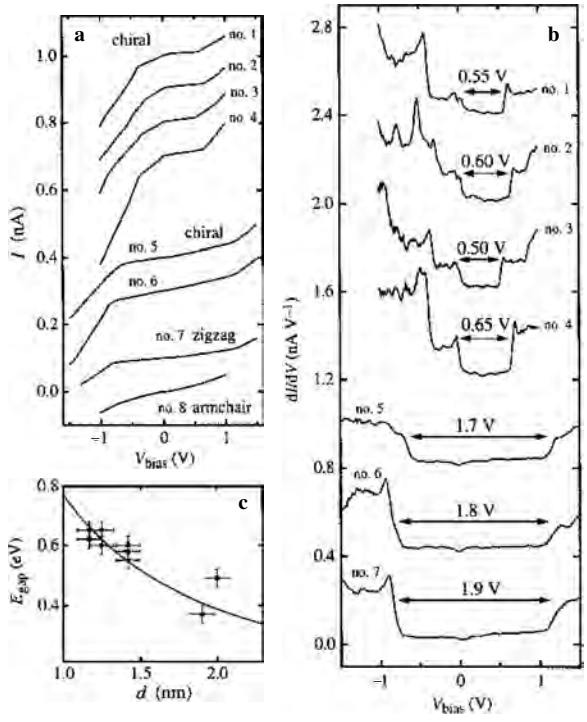


Figure 8. (a) Current–voltage curves obtained by tunneling spectroscopy on various nanotubes. (b) The derivatives *dI/dV* show two groups: a semiconducting one with gap values around 0.5–0.6 eV and a metallic one with gap values around 1.7–1.9 eV. (c) Energy gap versus diameter of semiconducting chiral tubes. Reprinted with permission from [65], J. W. G. Wildoer et al., *Nature* 391, 59 (1998). © 1998, Macmillan Magazines Ltd.

the dI/dV curves. There are two categories: the one has a well-defined gap values around 0.5–0.6 eV and the other has significantly larger gap values of ∼1.7–2.0 eV [65]. The gap value of the first category agrees very well with the predicted gap values for semiconducting tubes. As shown in Figure 9c, the energy gap decreases as the tube diameter d increases. This also agrees well with theoretical gap values obtained for an overlap energy $\gamma_0 = 2.7 \pm 0.1$ eV, which is close to the value $\gamma_0 = 2.5$ eV suggested for a single graphene sheet [3]. The very large bandgaps observed for the second category of tubes, 1.7–2.0 eV, are in good agreement with the values of 1.6–1.9 eV obtained from one-dimensional dispersion relations for a number of metallic tubes with $(n - m)/3$ being integer [65]. These metallic nanotubes are expected to have a small but finite DOS near the Fermi energy (E_F) and the apparent "gap" is associated with DOS peaks at the band edges of the next one-dimensional modes. Sharp van Hove singularities in the DOS are predicted at the onsets of the subsequent energy bands, reflecting the one-dimensional character of carbon nanotubes (see Fig. 7). The derivative spectra indeed show a number of peak structures (Fig. 8b). For semiconductors, it has been argued that $(dI/dV)/(I/V)$ represents the DOS better than the direct derivative *dI/dV*, partly because the normalization accounts for the voltage dependence of the tunnel barrier at high bias [68, 69]. In Figure 9, $(dI/dV)/(I/V)$ is shown, where sharp peaks are observed, resembling that predicted for van Hove singularities. The experimental peaks have a finite height and are broadened because of hybridization of wave functions. Raman scattering experiments also support the one-dimensional subband of nanotubes [70, 71]. Resistivity measurements of armchair SWNTs also suggested their metallic behavior, consistent with the theoretical calculation [72, 73]. In addition, momentum-dependent high-resolution electron energy-loss spectroscopy was performed on purified SWNTs [74]. Two groups of excitations have been found.

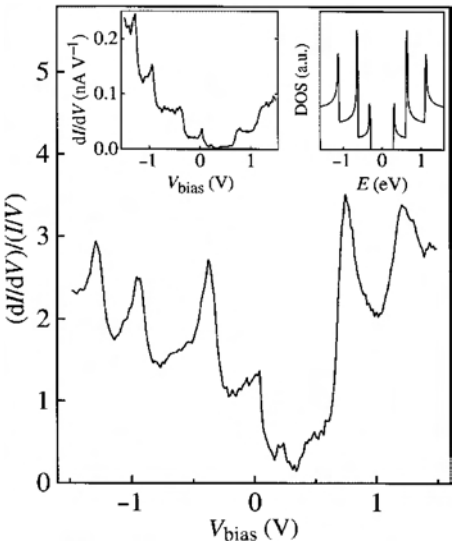


Figure 9. *(dI/dV)/(I/V)* which is a measure of the density of states versus V for nanotube 9. The left inset displays the raw *dI/dV* data, and the right inset is the calculated DOS. Reprinted with permission from [65], J. W. G. Wildoer et al., *Nature* 391, 59 (1998). © 1998, Macmillan Magazines Ltd.

One group is nondispersive and the energy position is characteristic of the separation of the van Hove singularities in the electronic DOS of the different types of nanotubes. The other one shows considerable dispersion and is related to a collective excitation of the π-electron system [74].

As pointed out by Dresselhaus [66], these results [66, 67] showed a wide range of helicities for SWNTs which contradict the narrow distribution of chiral angles determined by Raman scattering experiments [71] carried out on SWNT ropes synthesized by the same technique. In addition, van Hove singularities (energy threshold for an electronic subband) were also clearly resolved [65] in the experimental DOS of both chiral and achiral nanotubes. The observation of these well-separated and clearly resolved sharp spikes in the DOS of SWNTs with chiral symmetry was not expected [66]. Further calculation of the $\sigma + \pi$ electron densities of states of chiral carbon nanotubes using a tight-binding Hamiltonian showed that the electronic structures of SWNTs with chiral symmetry are similar to the zigzag and armchair ones [73, 74]. Fluorescence has also been observed directly across the bandgap of semiconducting carbon nanotubes, supporting the theoretical calculation of energy band structure of carbon nanotubes [77]. Most SWNTs synthesized using the current technologies are ropes consisting of many individual SWNTs. The first-principle calculation of electronic band structure of close-packing of individual nanotubes (10,10) into a rope showed that a broken symmetry of the (10,10) tube caused by interactions between tubes in a rope induces a pseudogap of about 0.1 eV at the Fermi level [78]. This pseudogap strongly modifies many of the fundamental electronic properties of carbon nanotube ropes. Structures of molecular electronic devices ultimately depend on tuning the interactions between individual electronic states and controlling the detailed spatial structure of the electronic wave functions in the constituent molecules. It is amazing that the two-dimensional images of electronic wave functions in metallic SWNTs have been obtained using STS [79]. These measurements reveal spatial patterns that can be directly understood from the electronic structure of a single graphite sheet. This represents an elegant illustration of Bloch's theorem at the level of individual wave functions.

3.3. Quantum Transport of Carbon Nanotubes

The electrical transport experiments on individual tubes are highly preferred. The first measurements on individual nanotubes were carried out on MWNTs [80–84]. Langer et al. [82] reported on electrical resistance measurements of an individual MWNT down to a temperature of $T =$ 20 mK. The conductance exhibited a $\ln T$ dependence and saturated at low temperature. A magnetic field applied perpendicular to the tube axis increased the conductance and produced aperiodic fluctuations. Their data also support two-dimensional weak localization and universal conductance fluctuations in mesoscopic conductors. These early studies on MWNTs suggested defect scattering, diffusive electron motion, and localization with a characteristic length scale of only a few nanometers. In addition, the electrical properties of individual MWNTs have been shown to vary strongly from tube to tube.

3.3.1. Ballistic Transport

It came as a surprise when the first experiments on individual SWNTs showed that nanotubes could have delocalized wave functions and behave as true *quantum wires* [85, 86]. Electrical measurement indicates that conduction occurs through well separated, discrete electronic states that are quantum-mechanically coherent over long distance, at least >140 nm [86]. Theory predicts that the electrons flow ballistically through carbon nanotubes and that the conductance is quantized [87–90]. Quantized conductance results from the electronic waveguide properties of extremely fine wires and constrictions. When the length of the nanotube is less than the mean-free path of electrons, the electronic transport is ballistic (i.e., each transverse waveguide mode or conduction channel contributes $G_0 = 2e^2/h$ to the total conductance). Theoretical calculation indicates that conducting single shell nanotubes have two modes or two conduction channels [87–90]; this predicts that the conductance of a single-walled nanotube is $2G_0$ independent of diameter and length. Another important aspect of ballistic transport is that no energy is dissipated in the conductor and the Joule heat is dissipated at the contacts of metal and nanotubes. Conductance measurements on MWNTs revealed that only one conduction channel G_0 exists in MWNTs, which conduct current ballistically over a length of 4 micrometers [91]. Recently, quantized conductance has been observed in SWNTs [92], which has two conduction channels $2G_0$, in agreement with the theoretical calculation. Theoretical studies also suggests that conduction electrons in armchair nanotubes experience an effective disorder averaged over the tube's circumference, leading to electron-mean-free paths that increase with nanotube diameter [93]. This increase should result in exceptional ballistic transport properties and localization lengths of 10 μm. For (10,10) armchair nanotubes, the mean-free path of 7.5 μm is obtained [93].

The fundamental reason for ballistic transport of carbon nanotubes is their *perfect* symmetric and periodic structure. It was shown that defects introduced into the nanotubes serve as scattering centers [94], which destroys the perfect structure. Theoretical calculation also showed that the absence of backscattering was demonstrated for single impurity with long range potential in metallic tubes [95–97] and a stepwise reduction of the conductance was inferred from multiple scattering on a few lattice impurities [98, 99]. Therefore, chemically doped semiconducting SWNTs may behave as diffusive conductors with shorter mean-free paths. It has been reported experimentally that mean-free paths of SWNTs are lower than the ones of reported structurally equivalent metallic SWNTs [100]. The backscattering contribution to the conductivity has been demonstrated to be more significant for doped semiconducting systems [101].

3.3.2. Other Transport Properties

Zeeman Effect Tans et al. [86, 102] observed an excited state by applying a magnetic field perpendicular to the tube axis, which moved relative to the ground state at a rate corresponding to a g-factor of 2.0 ± 0.5, consistent with the expected free-electron Zeeman shift. Cobden et al. [103] later studied the spin state by applying a magnetic field along the tube axis of the nanotube rope. It is concluded that as

successive electrons are added, the ground state spin oscillates between S_0 and $S_0 + \frac{1}{2}$, where S_0 is most likely zero. This results in the even/odd nature of the Coulomb peaks, which is also manifested in the asymmetry of the current–voltage characteristics and the peak height [106]. It is suggested that the g-factor of the Zeeman split is 2.04 ± 0.05 [104].

Aharonov–Bohm Effect When electrons pass through a cylindrical electrical conductor aligned in the magnetic field, their wavelike nature manifests itself as a periodic oscillation in the electrical resistance as a function of the enclosed magnetic flux. This phenomenon reflects the dependence of the phase of the electron wave on the magnetic field known as the Aharonov–Bohm effect [105], which causes a phase difference, and hence interference, between partial waves encircling the conductor in opposite directions. Theoretical studies showed [86, 106] that upon applying a magnetic field along the tube axis, the electronic structure of a carbon nanotube drastically changes from a metal to a semiconductor or from a semiconductor to a metal during variation of magnetic flux ϕ. The energy dispersion without the spin-$\boldsymbol{B}$ interaction is periodic in ϕ, with a period ϕ_0, as a result of the Aharonov–Bohm effect. Magnetoresistance measurements were carried out on individual MWNTs, which exhibit pronounced resistance oscillations as a function of magnetic flux [107]. The oscillations are in good agreement with theoretical predications for the Aharonov–Bohm effect in a hollow conductor with a diameter equal to that of the outermost shell of the nanotubes. Significant electron–electron correlation has been observed in experiments [108]. Electrons entering the nanotube in a low magnetic field are observed to have all the same spin direction, indicating spin polarization of the nanotube. When the number of electrons is fixed, variation of an applied gate voltage can significantly change the electronic spectrum of the nanotube and can induce spin-flips [108].

Luttinger Liquid Electron transport in conductors is usually well described by Fermi-liquid theory, which assumes that energy states of electrons near the Fermi level E_F are not qualitatively altered by Coulomb interactions. In one-dimensional systems, however, even weak Coulomb interactions cause strong perturbations. The resulting system, known as Luttinger liquid, is predicted to be distinctly different from its two- or three-dimensional counterpart [109]. Coulomb interactions have been studied theoretically for SWNTs [110, 111] and MWNTs [112]. Long-range Coulomb forces convert an isolated (N,N) armchair carbon nanotube into a strong renormalized Luttinger liquid [110]. At high temperatures, anomalous temperature dependence for the interaction, resistivity due to impurities, and power-law dependence for the local tunneling density of states were found. At low temperatures, the nanotube exhibits spin-charge separation, signaling a departure from orthodox theory of Coulomb blockade. Experimental measurements of the conductance of bundles ("ropes") of SWNTs as a function of temperature and voltage confirmed these theoretical studies [113].

4. SINGLE-ELECTRON TRANSISTORS

4.1. Single-Wall Carbon Nanotubes

In 1997, Bockrath et al. [85] reported the first single-electron transport of a single bundle containing 60 single-wall carbon nanotubes (10,10) with a diameter of 1.4 nm at a temperature of 1.4 K. The device structure (Fig. 10, left inset) consists of a single nanotube rope and lithographically defined Au electrodes. The device has four contacts and allows different segments of the nanotube to be measured. The device was mounted on a standard chip carrier and contacts were wire bonded. A dc bias (V_g) was applied to the chip carrier base to which the sample was attached. This dc bias can be used as gate voltage V_g to modify the charge density along the tube. Figure 10 shows the I–V characteristics of the nanotube section between contacts 2 and 3 as a function of temperature T. The conductance is strongly suppressed near $V = 0$ for $T < 10$ K. Figure 11A shows conductance G versus gate voltage V_g at $T = 1.3$ K. The conductance curve consists of a series of sharp peaks separated by regions of very low conductance. The peak spacing varies significantly. The height of peaks also varies widely with the maximum peak reaching e^2/h, where h is the Planck constant. The peak amplitude decreases with T (Fig. 11B) while the peak width increases with T (Fig. 11C). These phenomena can be understood based on Coulomb blockade effect described in Section 2.3. In this device, transport occurs by tunneling through the isolated segment of the rope. Tunneling on or off this segment is governed by the single-electron addition. The period of the peaks in gate voltage, ΔV_g, is determined by the energy for adding an additional electron to the rope segment.

In the same year (1997), Tans and co-workers [86] at Delft University of Technology built molecular devices using a metallic (armchair) SWNT as a quantum wire. Figure 12 shows the structure of the device. An individual SWNT with a diameter of ~1 nm is lying across two Pt electrodes with a separation of 140 nm. The third electrode located ~450 nm

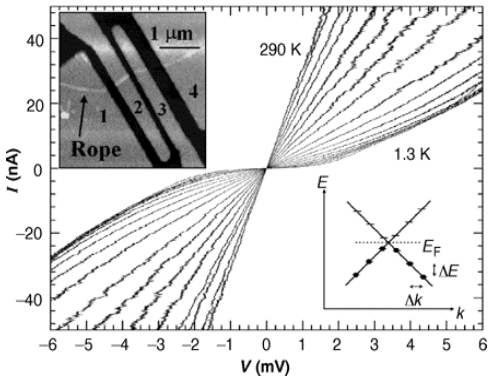


Figure 10. The I–V characteristics at a series of different temperatures for the rope segments between contacts 2 and 3. Left inset: AFM image of the fabricated device where the bright regions are metallic contacts. Right inset: Schematic energy-level diagram of the two 1D subbands near one of the two Dirac points with the quantized energy levels indicated. Reprinted with permission from [85], M. Bockrath et al., *Science* 275, 1922 (1997). © 1997, American Association for the Advancement of Science.

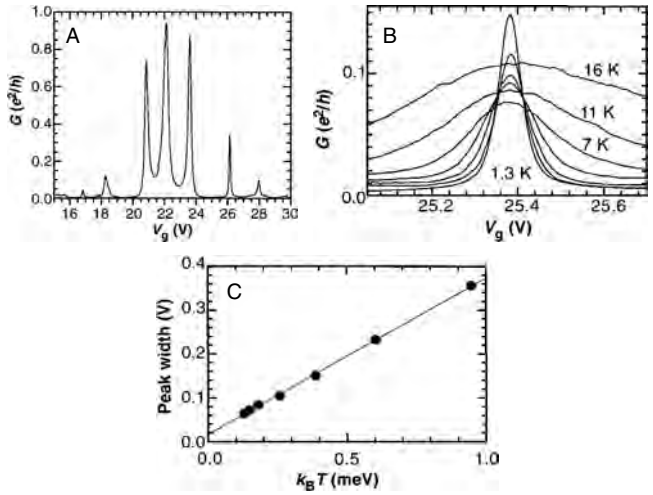


Figure 11. (A) Conductance G versus gate voltage V_g at $T = 1.3$ K for the rope segments 2 and 3. (B) Temperature dependence of a peak. (C) Width of the peak in (B) as a function of T. Reprinted with permission from [85], M. Bockrath et al., *Science* 275, 1922 (1997). © 1997, American Association for the Advancement of Science.

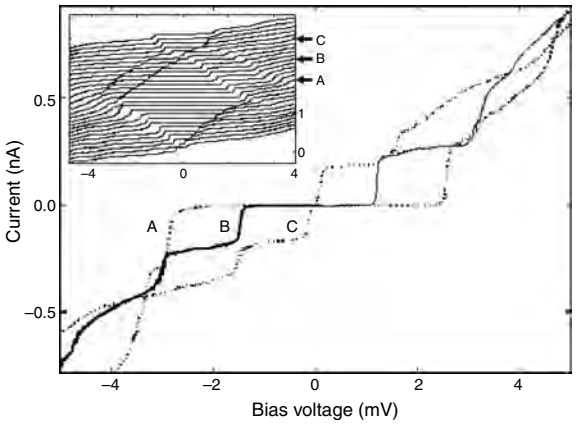


Figure 13. (a) Current–voltage curves of the nanotube at a gate voltage of 88.2 mV (trace A), 104.1 mV (trace B), and 120 mV (trace C). Inset: more I–V_{bias} curves with V_{gate} ranging from 50 mV (bottom curve) to 136 mV (top curve). (b) Current versus gate voltages at $V_{\text{bias}} = 30$ μV. Reprinted with permission from [86], S. J. Tans et al., *Nature* 386, 474 (1997). © 1997, Macmillan Magazines Ltd.

away from the nanotube functions as a gate. Their original idea is to build a single-electron transistor using a SWNT as a quantum wire. The electrical measurement was carried out at a low temperature of 5 mK. The typical current–voltage characteristics at various gate voltages are shown in Figure 13a. The coulomb charging effect is clearly observed. Coulomb charging occurs when the charging energy $E_c = (e^2/2C) \gg kT$. At low temperature, the Coulomb blockade effect can be observed. The two traces in Figure 13b were taken under identical conditions and show an occasional doubling of certain peaks. This bistability was regarded as the result of switching offset charges that shift the potential of the tube [86].

Chemical doping was used to achieve quantum dots and junctions for single-electron transistors [114]. Electrical measurements of the potassium (K) doped nanotube reveal single-electron charging at temperature up to 160 K [114]. The quantum dot is formed by inhomogeneous doping along the nanotube length [115–119]. The p–n–p junction

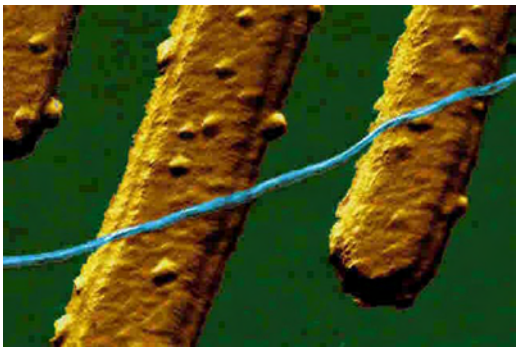


Figure 12. AFM tapping-mode image of a single-wall carbon nanotube on top of a Si/SiO_2 substrate with two 15-nm-thick Pt electrodes. Reprinted with permission from [86], S. J. Tans et al., *Nature* 386, 474 (1997). © 1997, Macmillan Magazines Ltd.

was obtained by chemical doping. The transport measurements of the junction showed that a well defined and highly reproducible on-tube single-electron transistor has been achieved [115]. It has been found that strong bends ("buckles") within metallic carbon nanotubes [2] act as nanometer-sized tunnel barriers for electron transport [116]. Single-electron transistors operating at room temperature have been fabricated by inducing two buckles in series within an individual metallic SWNT by manipulation with an AFM [117, 118]. The island with a length of 25 nm has been achieved and the resulting SET clearly showed the Coulomb blockade effect at room temperature [118]. Room temperature SETs have also been fabricated from SWNTs using V_2O_5 nanowires as masks for selective chemical doping [118]. Single-electron devices based on SWNTs with the line-shaped top gates [120], triple-barrier quantum dots [121], suspended quantum dots [122], field-induced p-type quantum dots [123], and kink-induced quantum dots [124] were fabricated. The microwave response of coupled quantum dots in SWNTs has also been measured [125]. The Coulomb oscillations for different microwave power were similar to those for different bias voltages without microwave.

Collins and co-workers [126] investigated electrical transport by sliding the STM tip along a nanotube. Figure 14 shows the schematic procedure for measuring nanotube characteristics using a single STM tip. From a position of stable tunneling (Fig. 14A), the STM tip was initially driven forward ~100 nm into the nanotube film (Fig. 14B). After retraction of the tip well beyond the normal tunneling range, nanotube material remained in electrical contact with the tip (Fig. 14C). Conductivity measurements were carried out by sliding the STM tip down along the nanotube while the tip remained electrically connected with the nanotube (Fig. 14D). The continuous motion of the tip allowed electrical characterization of different lengths of the nanotube. This technique results in a position-dependent electrical transport measurement along the extended lengths of selected nanotubes. A series of I–V curves were recorded

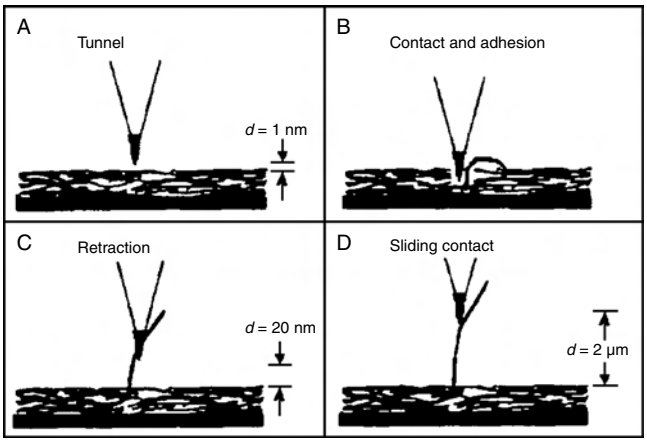


Figure 14. Schematic of the procedure for measuring nanotube characteristics with a single STM tip. Reprinted with permission from [126], P. G. Collins et al., *Science* 278, 100 (1997). © 1997, American Association for the Advancement of Science.

at positions 1600, 1850, 1950, and 2000 nm as shown in Figure 15. The first three curves are nonlinear but nearly symmetric. At a position of 2000 nm the I–V characteristics abruptly changed to a marked rectifying behavior (Fig. 15D). This response (Fig. 15D) was reproducible and persistent for positions up to 2300 nm before the nanotube was broken. It was suggested by Collins et al. [126] that the position-dependent behavior gives strong evidence for the existence of localized, well-defined, on-tube "nanodevices" with response characteristics consistent with the theoretical predictions. The extreme changes in conductivity were caused by contact with the localized nanotube defects that greatly altered the local $N(E)$. Although the injected current predominantly indicates a graphitic behavior for the nanotube rope, a nanotube defect at the contact point would obscure and dominate the transport characteristics. For example, the existence of a pentagon–heptagon defect in the otherwise perfectly hexagonal nanotube wall fabric can lead to sharp discontinuities in the electronic density state along the tube axis. It is possible to have one portion of the nanotube with metallic characteristics almost seamlessly joined to another portion that is semiconducting. This "junction" constitutes a pure-carbon Schottky barrier. The sliding STM probe indicates exactly this type of behavior as its position moves along the length of a nanotube by only a few nanometers, indicating the existence of a localized nanotube nanodevice.

4.2. Multiwall Carbon Nanotubes

Although single-electron transistors were made first from SWNTs [85, 86], a few reports [127–132] can be found for fabrication of SETs using MWNTs. Roschier et al. [127] of Helsinki University fabricated single-electron transistors using MWNTs through manipulation by a SPM. Figure 16 shows the experimental procedure for rotating and moving a nanotube, and eventually the tube was set across the electrodes with a gap of ~300 nm. The electrical measurements of the device were done at low temperatures

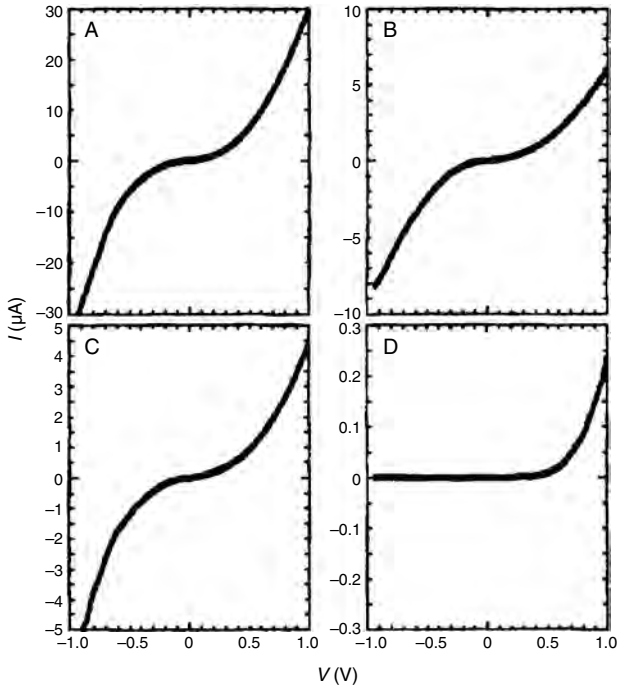


Figure 15. Different types of current–voltage characteristics, obtained for contact points at different heights of (A) 1600, (B) 1850, (C) 1950, and (D) 2000 nm along the carbon nanotube. Reprinted with permission from [126], P. G. Collins et al., *Science* 278, 100 (1997). © 1997, American Association for the Advancement of Science.

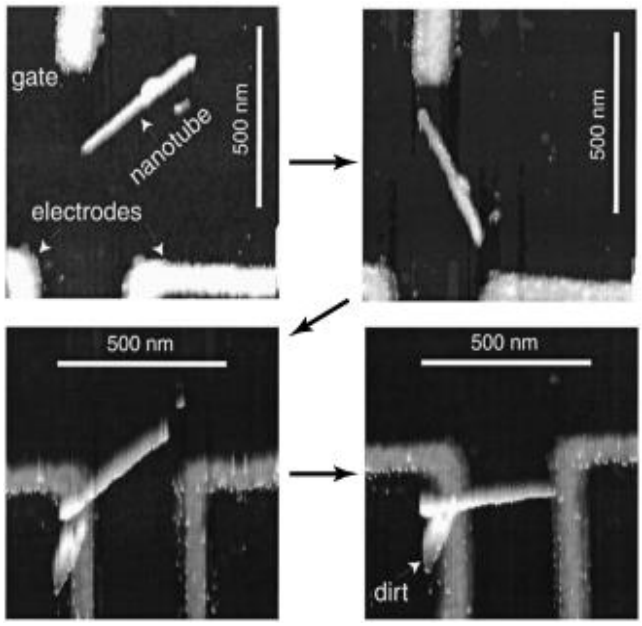


Figure 16. AFM images during moving process. The 410 nm long MWNT, the side gate, and the electrode structure are marked in the first frame. The last frame represents the measured configuration, where one end of the MWNT is well over the left electrode and the other end is lightly touching the right electrode. Reprinted with permission [128], L. Roschier et al., *Appl. Phys. Lett.* 75, 728 (1999). © 1999, American Institute of Physics.

down to 4 K. The measured I–V curves display a 15 mV wide zero current plateau across zero-voltage bias as shown in Figure 17. The Coulomb blockade effect is clearly observed below a few Kelvin and the nanotube behaves as a SET. The asymmetry of the gate modulation, illustrated in the inset for $V_{\text{bias}} = 10$ mV, indicates a substantial difference in the resistance of the tunnel junctions. There is a clear hysteresis in the I–V_{bias} curve at $T = 120$ mK. It is suggested that this phenomenon can be attributed to charge trapping, in which single electrons tunnel hysteretically across the concentric tubes. Roshier et al. [128] later constructed low-noise radio-frequency (rf) single-electron transistors using MWNTs. Contact resistance between a metal and a nanotube is commonly on the order of quantum resistance $R_Q = h/e^2 = 26.6$ kΩ. Hence, quantum fluctuations do not destroy charge quantization and thus it is possible to construct sensitive electrometers based on electrostatically controlled single-electron tunneling. The rf-SETs are the best electrometers at present [133]. As reported by Schoelkopf et al. [133] of Yale University, the sensitivity of rf-SETs based on Al islands approaches $1.2 \times 10^{-5} e/\sqrt{\text{Hz}}$, near the quantum limit at high frequencies. However, at frequencies below 1 kHz, these devices are plagued by the presence of $1/f^{\alpha}$ noise ($\alpha \sim$1–2). The origin of $1/f^{\alpha}$ noise is the trapping and detrapping of charges either in the vicinity of the island or on the surface of the nanotube or in the tunnel barrier [134, 135]. One way to reduce the $1/f^{\alpha}$ noise in SETs is to avoid contact of the central island with any dielectric material. In research by Roshier and co-workers [128], a freestanding MWNT across two electrodes was used as the island. The MWNT was moved onto the top of the electrodes by an AFM tip. The $1/f^{\alpha}$ noise of the SET is $6 \times 10^{-6} e/\sqrt{\text{Hz}}$ at 45 Hz, close to the performance in the best metallic SETs.

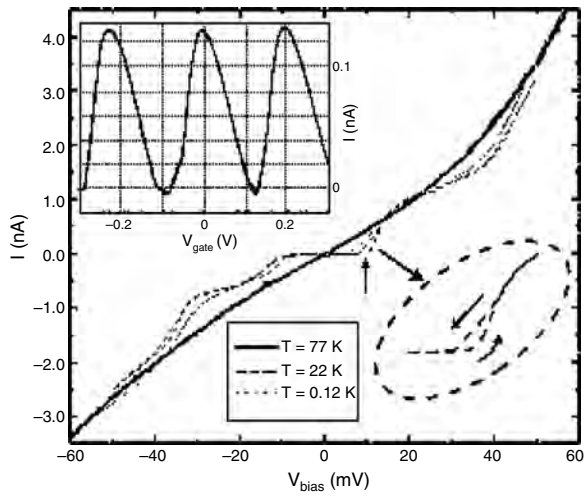


Figure 17. Measured I–V_{bias} curves at different temperatures when the gate is at zero bias. The inset shows the gate modulation at $V_{\text{bias}} = 10$ mV (indicated by the arrow) at $T = 120$ mK. The enlargement in the lower right corner shows the hysteretic behavior of the current. Reprinted with permission from [128], L. Roschier et al., *Appl. Phys. Lett.* 75, 728 (1999). © 1999, American Institute of Physics.

5. FIELD-EFFECT TRANSISTORS, LOGIC GATES, AND MEMORY DEVICES

5.1. Field-Effect Transistors

The significance of the paper by Tans et al. [86] is not in the quantum effect, but in the gate-induced modulation of conductance of the metallic nanotube. Field-effect transistors were first demonstrated using a single semiconducting SWNT by Tans et al. [136], and using both a SWNT and a MWNT by Avouris et al. [137–140]. Figure 18 shows the structure of the carbon nanotube field-effect transistor (CNTFET) [137]. The two electrodes are separated by 300 nm and gate oxide (SiO_2) is 140 nm. Figure 19 shows output characteristics of a SWNT-FET consisting of a single SWNT with a diameter of 1.6 nm for several values of the gate voltage. It is clearly seen that the source–drain current is modulated by electric field. The field effect of the MWNT-FET device was not observed [137]. The hole mobility is estimated to be 20 $\text{cm}^2/\text{V s}$. In 1999, Dai and co-workers [141] reported fabrication of FET using SWNTs controllably grown on substrates. Figure 20 shows the I–V curves at various gate voltages [141]. The asymmetry of the I–V curves was regarded as being inherent to the metal–tube–metal system. I–V curves after exchanging the source and drain show nearly unchanged asymmetry. These results suggest that the observed asymmetry is not caused by asymmetrical parameters such as different contact resistance at the two metal–tube contacts [141]. It was suggested that the asymmetry of I–V curves is due to high source–drain bias [141]. The transconductance was estimated to be 0.1 mS/μm.

5.1.1. Scaling of CNTFET

Theoretical studies [142] showed that the performance can be significantly improved if the channel length and gate oxide can be further scaled down. The I–V characteristics are similar to the ballistic Si MOSFETs except for the occurrence of quantized channel conductance. Because of ballistic transport, the average carrier velocity reaches 2.7×10^7 cm/s [145]. Theoretical studies [143] also show that the CNTFET can be scaled down to at least 5 nm. Because of the ballistic transport, there is no energy dissipation except at contacts, and terahertz operation may be possible. Recently Wind and co-workers at IBM [144] improved their CNTFET

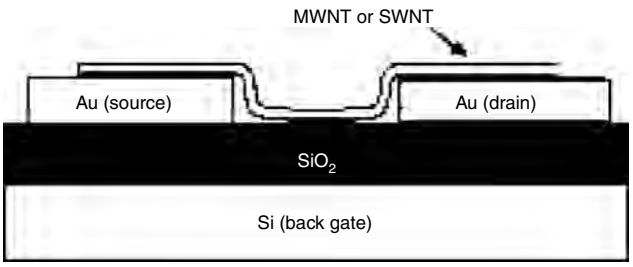


Figure 18. Schematic cross-section of the FET devices. A single nanotube of either multiwall or single-wall type bridges the gap between two gold electrodes. The silicon substrate is used as back gate. Reprinted with permission from [137], R. Martel et al., *Appl. Phys. Lett.* 73, 2447 (1998). © 1998, American Institute of Physics.

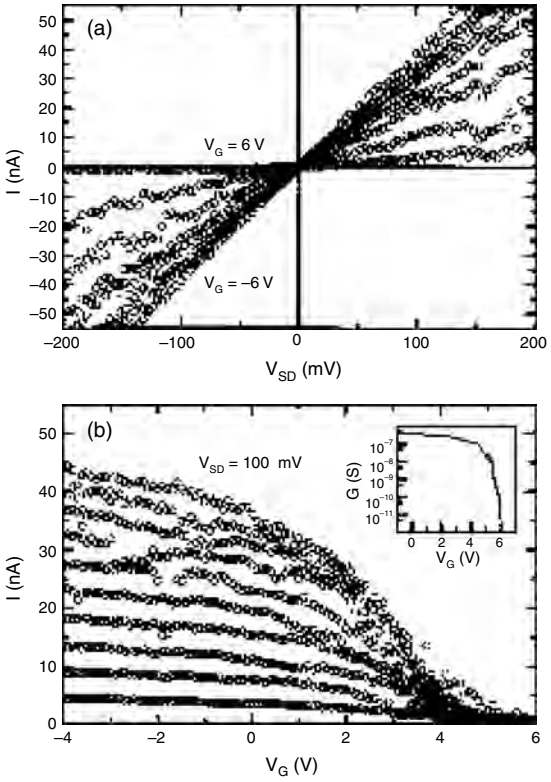


Figure 19. Output and transfer characteristics of a SWNT-FET: (a) I–V_{SD} curves measured for $V_G = -6$, 0, 1, 2, 3, 4, 5, and 6 V. (b) I–V_G curves for $V_{SD} = 10$–100 mV in steps of 10 mV. The inset shows that the gate modulates the conductance by 5 orders of magnitude ($V_{SD} = 10$ mV). Reprinted with permission from [137], R. Martel et al., *Appl. Phys. Lett.* 73, 2447 (1998). © 1998, American Institute of Physics.

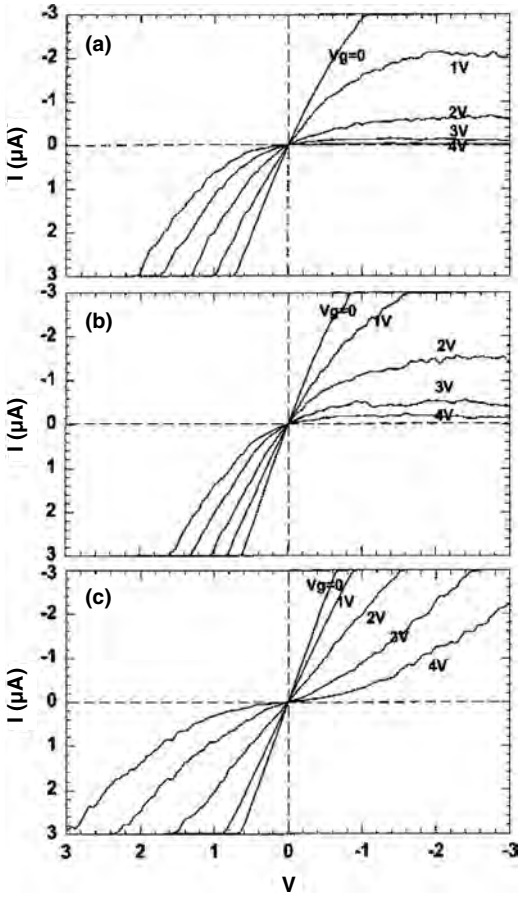


Figure 20. (a) Room-temperature I–V curves recorded with sample S1 for V in the range 3 to −3 V under various gate voltages. (b) I–V curves recorded after exchanging the source–drain electrodes. (c) Symmetrical I–V curves obtained by scanning V while biasing the two electrodes at $-V/2$ and $V/2$, respectively. Reprinted with permission from [141], H. Dai et al., *J. Phys. Chem. B* 103, 11246 (1999). © 1999, American Chemical Society.

structure with top gate and very thin gate oxide (15 nm). Figure 21 shows the device structure and its output characteristics. A single-wall carbon nanotube with a diameter of 1.4 nm was used as a semiconductor nanowire. The source and drain were defined by e-beam lithography with a gate length of 260 nm. I–V curves show excellent saturation and on–off ratio of 10^5. Table 1 shows a comparison of key device performance parameters for a 260 nm gate length p-type CNTFET with those of state-of-the-art Si MOS transistors, a 15 nm gate Si p-type MOSFET [145] and a 50 nm gate SOI p-type MOSFET [146]. It can be found that a CNTFET has superior performance over Si MOSFETs. A CNTFET exhibits a much higher ON current ($I_{on} = 2100\ \mu A/\mu m$), reasonable OFF current ($I_{off} = 150$ nA/μm), and very high transconductance (2321 μS/μm). It should be noted that the transconductance of the 15 nm gate Si p-MOSFET is only 975 μS/μm [145].

5.1.2. High Mobility

In [144], Wind et al. did not characterize the hole mobility during transport. I will analyze the hole mobility as follows. Because the I–V curves follow the classical transport model, the transconductance in saturation is expressed as [147]

$$G_{msat} = \frac{\mu_p C_{ox} W}{2L} |V_G - V_T|$$

where μ_p is the hole mobility. L and W are the gate length and gate width separately. C_{ox} is the gate oxide capacitance. V_G is the gate voltage and V_T is threshold voltage (−0.5 V). The gate width is considered to be half of the perimeter of the CNT (diameter = 1.4 nm). Thus the mobility can be calculated using

$$\mu_p = \frac{G_{msat} 2L}{C_{ox} W |V_G - V_T|}$$

Based on the given data of the transistor structure, the hole mobility is 2018 $cm^2\,V^{-1}\,s^{-1}$. This is much larger than the ideal hole mobility in bulk Si (~400 $cm^2\,V^{-1}\,s^{-1}$) and 200 × higher than the hole mobility (12 $cm^2\,V^{-1}\,s^{-1}$) derived from the 15 nm gate p-Si MOSFET [145]. *These surprising data indicate the potential of carbon nanotubes for high-speed device application similar to III–V compound semiconductors such as GaAs.* It was reported that SWNTs are extremely pure systems with large Fermi velocities of $v_F = 10^6$ m/s and ballistic transport over long distance [65, 91, 148]. Considering its unique 1D quantum wire electronic band structure and ballistic transport over long distance, it is highly possible for SWNTs to have extremely high mobility.

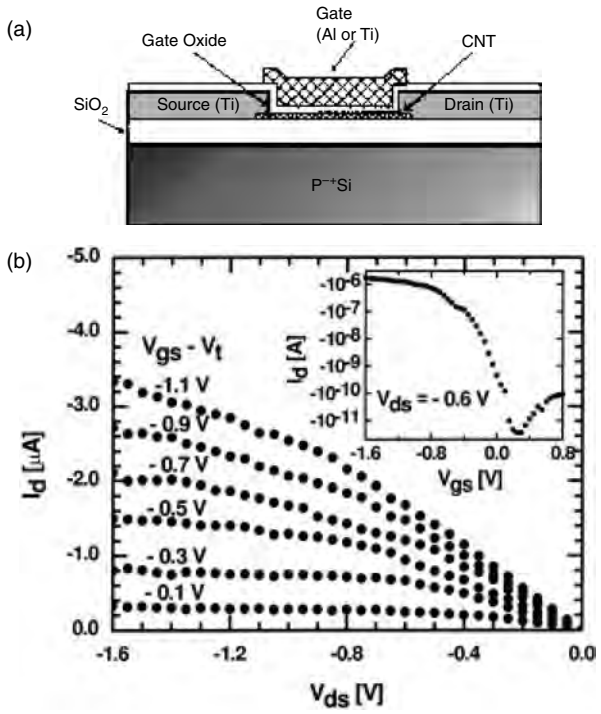


Figure 21. Schematic cross-section of top gate CNFET showing the gate and source and drain electrodes. (b) Output characteristic of a top gate *p*-type CNFET with a Ti gate and a gate oxide thickness of 15 nm. The gate voltage values range from −0.1 to −1.1 V above the threshold voltage, which is –0.5 V. Inset: Transfer characteristic of the CNFET for $V_{ds} = -0.6$ V. Reprinted with permission from [144], S. J. Wind et al., *Appl. Phys. Lett.* 80, 3817 (2002). © 2002, American Institute of Physics.

The high mobility of nanotube transistors estimated by the present author has been confirmed by Rosenblat et al. [149] and Kruger et al. [150]. Rosenblat et al. [149] constructed a carbon nanotube transistor using an electrolyte as gate, which was inspired by the study of doping effects using electrochemical gating [150, 151]. Figure 22 shows the device structure [149]. Catalyst islands containing $Fe(NO_3)_3 \cdot 9H_2O$, $MoO_2(acac)_2$, and alumina nanoparticles were defined on SiO_2. Carbon nanotubes were then grown by chemical vapor deposition. The source and drain with a separation of 1–3 μm (channel length) were defined using

Table 1. Comparison of key device performance parameters for a 260 nm gate length *p*-type CNTFET with those of state-of-the-art Si MOS transistors: a 15 nm-gate *p*-type Si MOSFET and a 50 nm gate *p*-type SOI MOSFET.

Types of transistors	CNTFET [144]	Si MOSFET [145]	SOI MOSFET [146]
Gate length (nm)	260	15	50
Gate oxide thickness (nm)	15	1.4	1.5
Threshold voltage (V)	−0.5	−0.1	−0.2
I_{ON} (μA/μm) @ $V_{DS} = V_{GS} - V_T = 1$ V	2100	265	650
I_{OFF} (nA/μm)	150	~500	9
Subthreshold slope (mV/dec)	130	~100	70
Transconductance (μS/μm)	2321	975	650

Source: Adapted with permission from [144], S. J. Wind et al., *Appl. Phys. Lett.* 80, 3817 (2002). © 2002, American Institute of Physics.

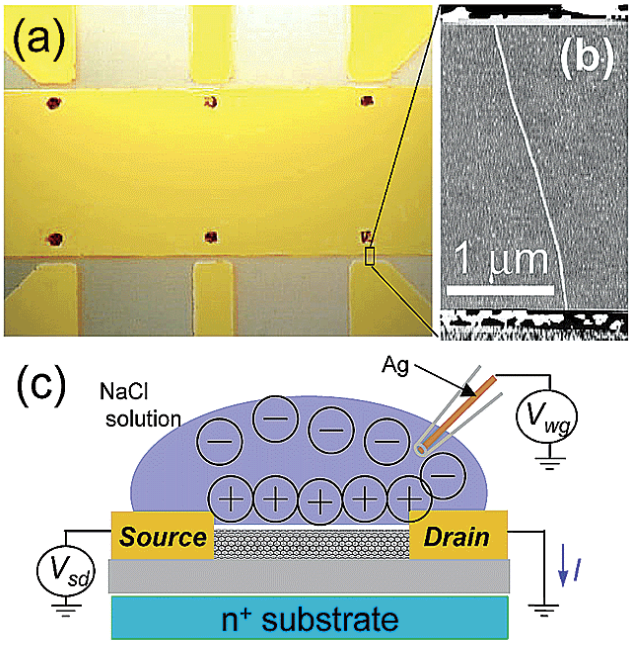


Figure 22. Optical micrograph of the device. Six catalyst pads (dark) can be seen inside the area of the common electrode. Correspondingly, there are six source electrodes for electrical connection to tubes. (b) AFM image of a tube between two electrodes. The tube diameter is 1.9 nm. (c) Schematic of the electrolyte gate measurement. A water gate voltage V_{wg} is applied to droplets through a silver wire. Reprinted with permission from [149], S. Rosenblat et al., *Nano Lett.* 2, 869 (2002). © 2002, American Chemical Society.

photolithography and a lift-off process. A micropipet is used to place a small (~10–20 μm) saltwater droplet (NaCl solution) over the nanotube device. A voltage V_{wg} applied to a silver wire in the pipet is used to establish the electrochemical potential in the electrolyte relative to the device. Then the electrolyte functions as a liquid gate. The output characteristics of the electrolyte carbon nanotube FET are similar to those in Figure 21. The transconductance of the transistor reaches its maximum ~20 μS at a gate voltage of −0.8 V. The mobility inferred from the conductance measurement is in the range of 1000 to 4000 $cm^2\,V^{-1}\,s^{-1}$. The maximum on-state conductance is also shown for the same samples. Values on the order of e^2/h are routinely obtained, within a factor of 4 of the theoretical limit of $4e^2/h$. Fuhrer et al. [152] reported the hole mobility in a SWNT of 9000 $cm^2\,V^{-1}\,s^{-1}$, which is the highest value ever reported.

In addition, the field effect is clearly shown even at a temperature of 5 K with a spikelike conductance, which is attributed to the van Hove singularities [153]. AFM tips were used to apply pointlike local gates to manipulate the electrical properties of an individual SWNT contacted by Ti electrodes [154]. The AFM tip contacting on a semiconducting SWNT causes depletion at a local point, leading to orders of magnitude decrease of the nanotube conductance, while local gating to a metallic SWNT leads to no change in conductance [154]. Theoretical study also suggests that a quantum dot is formed because of the induced *n*-type together with the *p*-type near the metal contacts when a positive gate voltage is applied on the *p*-type SWNT [155]. The induced quantum dot enhances the conductance.

Because the energy bandgaps of semiconducting SWNTs are inversely proportional to their diameter, large-diameter SWNTs have smaller energy bandgaps. Transistors made of large-diameter SWNTs exhibit ambipolar field-effect transistor behavior [156, 157]. Theoretical study showed that the energy gap of a semiconducting nanotube can be narrowed, when the tube is placed in an electric field perpendicular to the tube axis (e.g., in the FET case) [158]. This band-structure modulation may affect the electrical properties of CNTFETs. No experimental research has been reported regarding this phenomenon. For characterization of the semiconductor/oxide interface, capacitance–voltage measurement is usually carried out on MOS capacitors. Theoretical study showed that the calculated C–V curves reflect the local peaks of the 1D DOS in the nanotube [159]. This might be used for diagnose the electronic structure of nanotubes, providing a more convenient approach than STM. However, experimental measurement of the capacitance in nanoscale is not easy because the accumulation capacitance of a 1-μm long nanotube MOS capacitor is only 1.5×10^{-4} pf or 1.5 pf/cm [159].

5.2. Logic Gates

In 2001, several groups [160–162] demonstrated logic circuits using carbon nanotube transistors. Bachtold et al. [160] showed inverter, NOR gate, static random access memory (SRAM), and ring oscillator. Derycke et al. [161] and Liu et al. [162] showed the CMOS inverter using both n- and p-channel CNTFETs. Figure 23 shows individual device structure and layout [160]. Unlike the previous nanotube transistor structure using back gate, which applies the same gate voltage to all transistors, the transistor structure consists of a microfabricated Al wire with native Al_2O_3 as gate insulator, which lies beneath a semiconducting nanotube that is electrically contacted to two Au electrodes. The channel length is about 100 nm and gate oxide thickness is about a few nanometers. This layout allows the integration of multiple nanotube FETs. The transistor is a p-type enhancement mode FET with transconductance of 0.3 μS and on–off ratio of at least 10^5. The transistor can operate at an ON current of 100 nA and an OFF current of 1 pA. The basic logic elements such as inverter, NOR gate, SRAM, and ring oscillator were constructed as shown in Figure 24. The inverter exhibits very good transfer characteristics. When input voltage is -1.5 V (logic 1), the output voltage is 0 V (logic 0). When the input voltage is switched to 0 (logic 0), the output becomes -1.5 (logic 1). Although the transition is not as sharp as a Si MOSFET, it is still competitive. Because this inverter is constructed using a single transistor, the standby current is still high. The ring oscillator shows good oscillation waveforms although the oscillation frequency is low in this pioneer stage. The inverters have been constructed using complementary nanotube FETs similar to Si CMOS structure (complementary MOS), leading to minimum standby power consumption [161, 162]. Figure 25 shows the CMOS inverter based on both n- and p-CNTFETs and its transfer characteristic [161]. A single nanotube bundle is positioned over the gold electrodes to produce two p-type CNTFETs in series. The device is covered by PMMA and a window is opened by e-beam lithography to expose part of the nanotube. Potassium is then evaporated through this window to produce an n-CNTFET, while the other CNTFET remains p-type. The transfer characteristics show much better transition region (more steep slope).

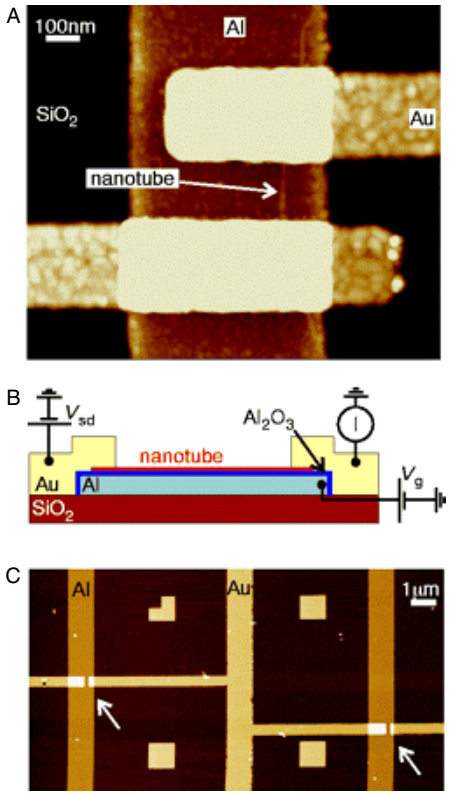


Figure 23. Device layout. (A) Height image of a single-nanotube transistor, acquired with an atomic force microscope. (B) Schematic side view of the device. (C) Height-mode atomic force microscope image of two nanotube transistors connected by a Au interconnect wire. The arrows indicate the position of the transistors. Reprinted with permission from [160], A. Bachtold et al., *Science* 294, 1317 (2001). © 2001, American Association for the Advancement of Science.

5.3. Memory Devices

A concept for molecular electronics exploiting carbon nanotubes as both molecular device elements and molecular wires for reading and writing information has been proposed [163]. Each device is based on suspended, crossed nanotube geometry that leads to bistable, electrostatically switchable ON/OFF states. The device elements are naturally addressable in large arrays by the carbon nanotube molecular wires making up the devices. These reversible, bistable device elements could be used to construct nonvolatile random access memory and logic function tables at an integration level approaching 10^{12} elements per cm^2, and an element operation frequency in excess of 100 GHz [163]. However, strictly speaking, these memory devices or logic gates are not made of CNTFETs. Several groups [152, 164–166] reported fabrication of memory devices using nanotube field-effect transistors. Air-stable n-type, ambipolar CNTFETs were fabricated and used in nanoscale memory cells [164]. The n-type transistors are achieved by annealing nanotubes in hydrogen gas and contacting them by cobalt electrodes. Due to their nanoscale capacitance, CNTFETs

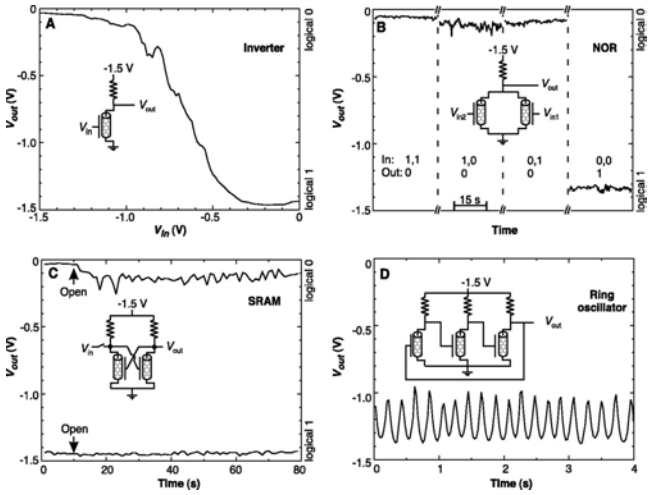


Figure 24. Demonstration of one-, two-, and three-transistor logic circuits with carbon nanotube FETs. (A) Output voltage as a function of the input voltage of a nanotube inverter. Inset: Schematic of the electronic circuit. The resistance is 100 MΩ. (B) Output voltage of a nanotube NOR for the four possible input states (1,1), (1,0), (0,1), and (0,0). The resistance is 50 MΩ. (C) Output voltage of a flip–flop memory cell (SRAM) composed of two nanotube FETs. The two resistances are 100 MΩ and 2 GΩ. (D) Output voltage as a function of time for a nanotube ring oscillator. The three resistances are 100 MΩ, 100 MΩ, and 2 GΩ. Reprinted with permission from [160], A. Bachtold et al., *Science* 294, 1317 (2001). © 2001, American Association for the Advancement of Science.

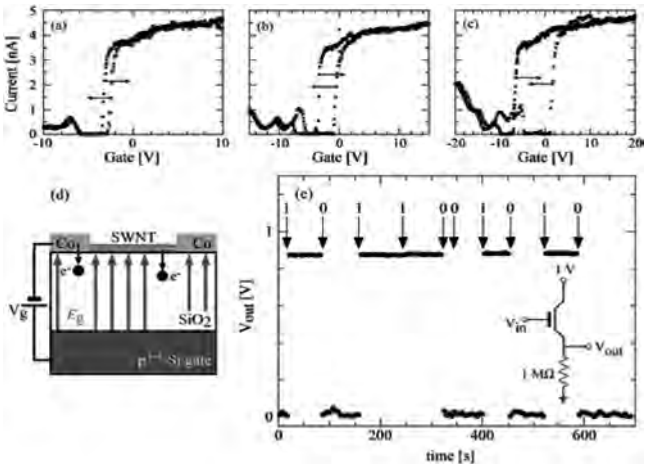


Figure 26. (a)–(c) High vacuum I–V_g data at $V_{ds} = 0.5$ mV. Device hysteresis increases steadily with increasing V_g due to avalanche charge injection into bulk oxide traps. (d) Diagram of avalanche injection of electrons into bulk oxide traps from the CNFET channel. (e) Data from CNFET-based nonvolatile molecular memory cell. A series of bits is written into the cell (see text) and the cell contents are continuously monitored as a voltage signal (V_{out}) in the circuit shown in the inset. Reprinted with permission from [164], M. Radosavljevic et al., *Nano Lett.* 2, 761 (2002). © 2002, American Chemical Society.

are extremely sensitive to the presence of individual charges around the channel, which can be used for data storage that operate at the few-electron level [165]. Figure 26 shows the threshold voltage shift due to storage of charges (a,b,c), device structure (d), and the voltage signal (V_{out}) due to charge storage [164]. In addition, the data-storage stability of over 12 days has been achieved [165].

6. DOPING, JUNCTIONS, AND METAL–NANOTUBE CONTACTS

A key technology advancement for the success of semiconductor industry is achievement of *n*- and *p*-type doping, junctions, and Ohmic contacts between metal and

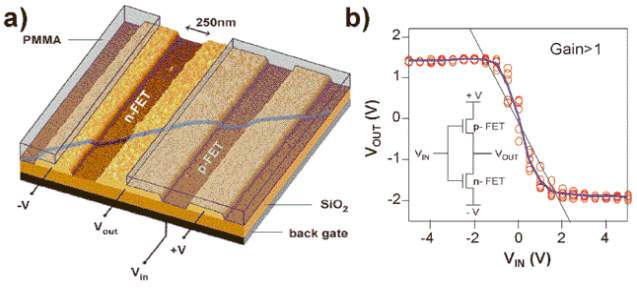


Figure 25. (a) AFM image showing an intramolecular logic gate. (b) Characteristics of the resulting intramolecular voltage inverter. The thin straight line corresponds to an output/input gain of one. Reprinted with permission from [161], V. Derycke et al., *Nano Lett.* 1, 453 (2001). © 2001, American Chemical Society.

semiconductor. It is critical to achieve doping, *pn* junctions, and Ohmic contacts for carbon nanotubes so that nanotube electronics may evolve into a large industry.

6.1. Chemical Doping

Antonov and Johnson [167] observed current rectification in a molecular diode consisting of a semiconducting SWNT and an impurity. It was suggested that rectification resulted from the local effect of the impurity on the tube's band structure. It is not clear what type of impurity it was. Lee et al. [168] reported doping of SWNTs by vapor-phase reactions with bromine and potassium. Doping decreases the resistivity of SWNTs at 300 K by up to a factor of 30 and enlarges the region where the temperature coefficient of resistance is positive, which is the signature of metallic behavior. It was reported [169, 170] that potassium (K) doping of SWNTs creates *n*-type carrier (electrons). The doping effects were studied using the transistor structure. The SWNT ropes were placed on top of Au electrodes that have 500 nm separation. The electrodes were fabricated on the oxidized n^+-Si substrate which serves as gate. The Au electrodes serve as source and drain. Figure 27 shows conductance vs gate voltage for an undoped nanotube rope (open circles) and an nanotube doped with potassium (solid circle) [169]. For the undoped nanotube, the conductance increases with decreasing gate voltage, indicating *p*-type behavior. For the K-doped nanotube, the conductance increases with increasing gate voltage, indicating *n*-type behavior. The typical values for the carrier density are found to be ∼100–1000 electrons/μm and the effective mobility of electrons is $\mu_{eff} \sim$ 20–60 in early time [169]. Derycke et al. [171] reported two methods for conversion of SWNTs from *p*-type to *n*-type. The first method involves conventional doping with an electron donor and the second

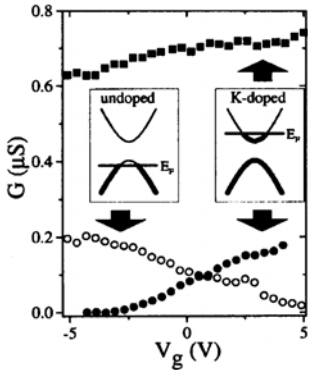


Figure 27. The conductance decreases with increasing gate voltage for an undoped sample (open circles), indicating *p*-type behavior. Left inset shows the energy band corresponding to this case. The conductance increases with increasing gate voltage after potassium doping at a high doping level (solid squares) and at a low doping level (solid circles). Right inset shows the energy band corresponding to this situation. Reprinted with permission from [169], M. Bockrath et al., *Phys. Rev. B* 61, R10606 (2000). © 2000, American Physical Society.

consists of annealing the metal nanotube contacts in vacuum to remove adsorbed oxygen. It has been found that the main effect of oxygen adsorption is not to dope the bulk of the nanotube, but to modify the Schottky barriers at the metal–semiconductor contacts [172, 173]. It also found that boron-doped MWNTs showed an enhanced room temperature conductivity [174]. However, theoretical calculation showed that H_2O adsorption on a SWNT reduces conductance [175].

6.2. Junctions and Metal–Nanotube Contacts

Zhou et al. [176] published the first attempt for controllable chemical doping of individual carbon nanotubes to achieve *pn* junctions. Figure 28 shows the device structure for potassium doping and the resulting energy band structure [176]. The SWNT used in the doping experiments has a diameter of ~2 nm and a length of 3.5 μm and is placed across two metal electrodes. A back gate is used to modulate the carrier concentration by applying voltage on the gate. A polymethylmethacrylate (PMMA) layer of 340 nm thickness covers the left half of the nanotube, leaving the right half exposed. Prior to doping, SWNT is a *p*-type semiconductor. Potassium doping of the SWNT is carried out in vacuum by electrical heating of a potassium source. Figure 29A shows the I–V_g curves [176]. When $V_g < -1$ V (Regime I), hole accumulation is achieved in the SWNT, leading to p^+ in the PMMA protected area due to increase in hole concentration and n in the exposed region due to compensation of electrons. In Regime II, electron concentration begins to increase, resulting in p^+n^+ junction. I–V curves of p^+n and p^+n^+ junctions are shown in Figure 29B and C. In the forward bias regime, a nice *pn* junction diode is demonstrated. However, the reverse bias breakdown voltage is too small to be qualified as a diode [176]. Anyway, this is the first result showing some possibility for fabrication of nanotube-based *pn* junction diodes.

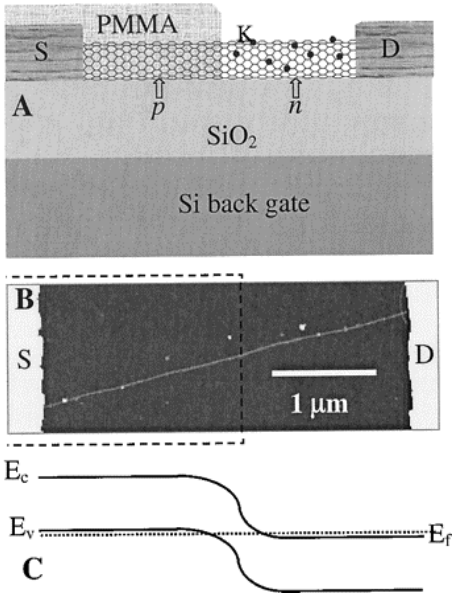


Figure 28. An SWNT with modulated chemical doping. (A) Device scheme with the SWNT contacted by two Ni/Au electrodes. The right half of the SWNT is doped by evaporating K atoms (black dots) from an alkaline metal at 10^{-6} Torr. (B) Atomic force microscopy image of the SWNT recorded before PMMA coating of the left half and doping. The bright regions at the two ends are the electrodes. Dashed lines are drawn to highlight the later PMMA-covered region. (C) A band diagram for the system. Reprinted with permission from [176], C. Zhou et al., *Science* 290, 1552 (2000). © 2000, American Association for the Advancement of Science.

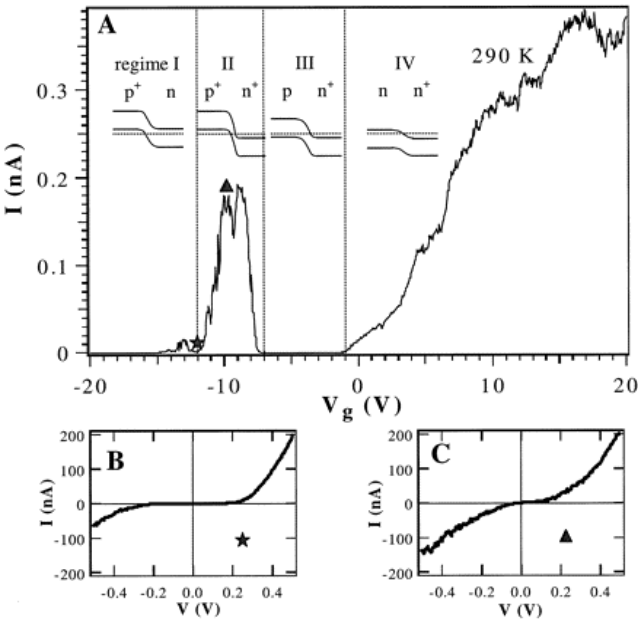


Figure 29. Electrical properties of the modulation-doped SWNT at room temperature. (A) Current versus gate-voltage I–V_g curve recorded under a bias voltage $V = 1$ mV. The drawings show the band diagrams in four regimes. (B) and (C) I–V curves recorded in Regimes I and II. The star and triangle in (A) marks the gate voltages for the I–V curves in (B) and (C), respectively. Reprinted with permission from [176], C. Zhou et al., *Science* 290, 1552 (2000). © 2000, American Association for the Advancement of Science.

Based on theoretical calculation, Chico et al. [177] proposed metal/semiconductor or semiconductor/semiconductor junctions, made of SWNTs, and based on the introduction of topological defects in nanotubes. By introducing a pentagon and a heptagon into the hexagonal carbon lattice, two tube segments with different electronic structures can be seamlessly fused together to create semiconductor–semiconductor or metal–semiconductor junctions [172, 178, 179]. Two carbon nanotubes have also been fused together by high electric field [180]. The CN_x/nanotube junctions have been synthesized by a microwave plasma assisted CVD method [181]. It is of particular interest that two SWNTs were crossed over each other to form junctions [182]. Theoretical study also suggests that negative differential resistance may be observed in metal–nanotube–metal structures [183]. Figure 30 shows the AFM image of a nanotube junction with a sharp kink of about 40° [172]. The kink consists of five to seven defects. Figure 31 shows the I–V characteristics of the kink metal–semiconductor junction [172]. The rectification effect is clearly seen in the figure. A model has been proposed for the kink-shaped carbon nanotube Schottky diode, where the gate voltage modulation is included [184, 185]. It has been observed that the nanotube Schottky diodes [172, 179] and the nanotube *pn* junctions [176] exhibit much lower reverse-bias breakdown voltage than conventional, micrometer-size Schottky diodes and *pn* junctions. A theoretical model has been proposed to describe the potential barrier shape in ultrasmall Schottky diodes [186]. It is suggested that for diodes smaller than a characteristic length l_c (e.g., 80 nm for $N_d = 10^{17}$ cm^{-3}), Schottky barrier thickness becomes a function of the diode size. Consequently, the contribution of tunneling to the total conductance is dramatically enhanced, resulting in lower reverse breakdown voltage in nanoscale diodes [186]. The carbon nanotube "T" junction with heptagons or pentagons for joints has been proposed theoretically [187]. However, experimentally no "T" junction has been observed. Instead, "Y" junctions have been produced through an anodic aluminum oxide (AAO) template [188, 189] and directly on substrates [190–196]. Theoretical calculation confirms the rectification

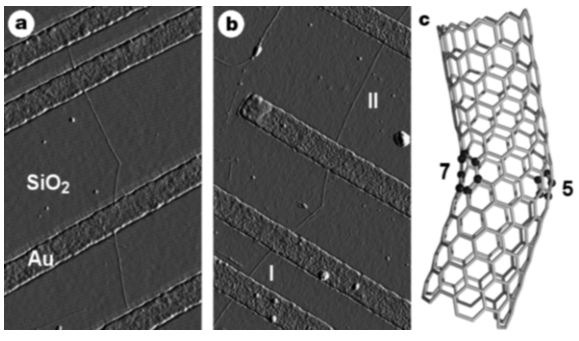


Figure 30. Tapping-mode atomic force microscope amplitude images of examples of nanotube junction devices. (a) and (b) Nanotubes that contain a single kink of 36° and 41° respectively. (c) Illustration of the carbon-bond network of a kink junction constructed between an "armchair" tube and a "zigzag" tube, where 5 denotes a pentagon, 7 denotes a heptagon, and the atoms in the pentagon and heptagon are highlighted by dark balls. Reprinted with permission from [172], Z. Yao et al., *Nature* 402, 273 (1999). © 1999, Macmillan Magazines Ltd.

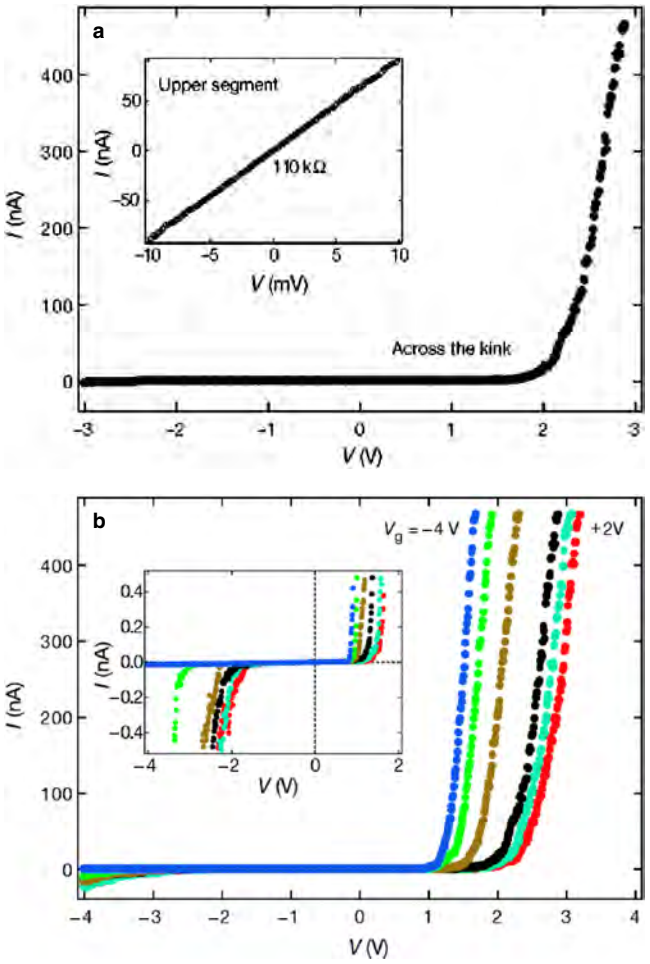


Figure 31. Current–voltage characteristics across the metal–semiconductor junction of (a), showing rectifying behavior. The data are taken at 100 K. The results at room temperature are similar, but the data are noisier. Inset in (a): the I–V curve for the upper straight segment measured at room temperature. In (a), the gate is grounded. In (b): the gate voltages from right to left are 2, 1, 0, −1, −2, and −4 V respectively. Inset in (b): expanded view of the small-current region which shows more clearly the onset of conduction for both bias polarities. Reprinted with permission from [172], Z. Yao et al., *Nature* 402, 273 (1999). © 1999, Macmillan Magazines Ltd.

effect of "Y" junctions [197–199]. When the positively biased metal–nanotube contact (source electrode) is locally gated with a negative gate voltage, a dramatic increase in transport current occurs because the Schottky barrier thickness is reduced due to accumulation of holes near the contact [200]. For the negatively biased metal–nanotube contacts, the gate voltage has no effect on transport because no Schottky barrier exists there [200]. Therefore, by positioning the gate locally near one of the contacts, the nanotube FET is converted into a rectifying diode [200].

Zhang et al. [201] fabricated the first real Ohmic contacts between SWNTs and Ti by solid solid reaction: C (nanotubes) + M (solid) → MC (solid), where M is metal. The reaction was performed at a temperature ranging from 800 to 1000 °C in ultrahigh vacuum or an inert atmosphere to avoid volatile reactant [201]. The continuous transformation of the SWNT to carbide is controlled by the diffusion of

M to the C/M interface. The I–V curves showed a straight line when the bias voltage is swept from −1 to 1 V, indicating a true *Ohmic contact* [201]. The resistance at the contact is ∼0.3 kΩ [201], much better than the contact resistance (∼20 kΩ) for the direct metal/nanotube contacts [202, 203].

7. NANOTUBE-BASED NANOFABRICATION

Carbon nanotubes exhibit Coulomb-blockade effects, ballistic transport, and field effects. Of particular importance is the field effect that may lead to the carbon nanotube electronics era. However, a critical issue for successful realization of carbon nanotube electronics is how to manipulate the nanotubes so that large-scale manufacturing is possible.

7.1. Manipulation of Nanotubes Using AFM and STM

Since discovery of carbon nanotubes [1], for electronic device research, the majority of research has been focused on manipulation of carbon nanotubes using the STM or AFM [86, 118, 127, 163, 203–223]. Tans et al. [86] deposited SWNTs on the SiO_2/Si substrate which had patterned metal electrodes. An individual nanotube wire was put across the Pt electrodes, which was imaged by an AFM. Physical manipulation of nanotubes with the AFM tip, like rolling, sliding, bending, and buckling, has been used to investigate mechanical properties of nanotubes [211, 212]. Roschier et al. [127] positioned a semiconducting multiwalled nanotube between two gold electrodes at the SiO_2 surface. The 410 nm long MWNT was manipulated and positioned on the gold electrodes by the AFM tip. Soh et al. [203] synthesized SWNTs on the patterned catalytic islands, which are contacted with metal pads. Lefebvre et al. [204] developed a method to assemble SWNT circuits using a tapping mode AFM. Nanotubes can be controllably translated, rotated, cut, and placed on top of one another by varying the tip–sample force and the tip speed. These operations can construct complex nanotube circuits. Of particular interest is that the suspended SWNT can be deformed locally by the AFM tip, leading to a decrease in conductance of the nanotube by two orders of magnitude [217–219]. This effect can be used to construct nanoelectromechanical devices. Tight-binding simulation showed that this effect is caused by the formation sp^3 bonds because of the mechanical pushing action of the tip [217, 218]. Collins et al. at IBM [205] demonstrated a simple method for permanently modifying MWNTs by using current-induced breakdown to eliminate individual shells one at a time. Carbon nanotubes can withstand remarkable current densities, exceeding 10^9 A/cm^2, because of their strong carbon–carbon bonding. However, at high enough current nanotubes ultimately fail. In MWNTs, this failure occurs in air at a certain threshold power through the rapid oxidation of the outermost carbon shell. The mechanism for the breakdown initiation is the current-induced defect formation. By using the electrical breakdown technique, they can remove the MWNT shells one by one. This controlled breakdown technique is also performed with the help of STM or AFM. Study of the reliability and current carrying capacity showed that under high current density ($>10^9$ A/cm^2) no observable failure in the nanotube structure and no measureable change in the resistance are detected at temperatures up to 250 °C and for time scales up to 2 weeks [224]. *Although these processes and techniques are especially useful for study of nanotubes' physical properties, they may not be appropriate for large-scale fabrication of nanotube-based devices and circuits with high yield.*

7.2. Controllable Growth and Placement of Nanotubes

Before we can think about building sophisticated nanotube-based circuitry, we must find how to grow the nanotubes in specific locations, orientations, shapes, and sizes, as well as how to construct nanotube devices at specific locations and how to connect nanotube devices with each other.

7.2.1. Vertically Aligned Carbon Nanotubes

Li et al. [225, 226] reported the first large-scale synthesis of vertically aligned carbon nanotubes by using CVD catalyzed by iron nanoparticles embedded in mesoporous silca. Vertically aligned arrays of isolated tubes with spacing between the tubes of about 100 nm were controllable through the pores. This method has been extended to grow freestanding carbon nanotubes on glass substrates at lower temperature (<700 °C) using nickel film as catalysts by CVD [227–229] and also on Si substrates [231]. Because the mesoporous structure of silica was not uniform, the controllability of tubes was not very good. It is well known that the porous structure of AAO film exhibits hexagonal close-packed nanopores with large-scale periodicity and high densities [232–234]. Highly ordered and vertically aligned carbon nanotube arrays was obtained when growth of nanotubes was carried out on AAO porous structure with cobalt particles deposited into the bottom of the pores [235–241]. Figure 32 shows the procedure for fabrication and growth and the scanning electron microscope (SEM) image of arrays of nanotubes [236]. It should be noted that these highly ordered CNT arrays were successfully fabricated on aluminum substrates. In order to incorporate the carbon nanotubes into the Si electronics, growth of aligned nanotubes on Si substrates is imperative. AAO films were cracked at about 300–400 °C, whereas AAO on N_b was not damaged at 1100 °C in He gas because the thermal expansion coefficient of N_b is close to that of AAO. It is suggested that the N_b layer under AAO may afford the durability for the high temperature process [242]. CNTs were grown on the AAO film based on 100-nm N_b film [242]. Unfortunately, the as-grown nanotubes were not uniformly controllable and sparse nanotubes were grown [242]. Hu et al. at the University of Kentucky [243, 244] successfully fabricated highly ordered CNT arrays on *silicon substrate* using AAO template by the flame synthesis technique (see Fig. 33). In the process, the Al film was not completely consumed during anodization and the Al film that remained between the AAO and Si substrate served as a buffer layer, resulting in smooth surface without cracks after the high temperature process. This demonstrated the feasibility for integration of carbon nanotube devices with Si integrated circuits.

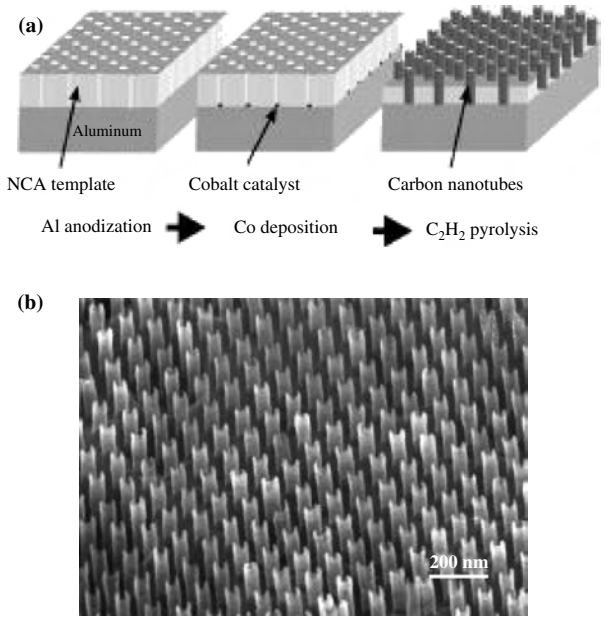


Figure 32. Schematic of fabrication process. (b) SEM image of the resulting hexagonally ordered array of carbon nanotubes fabricated using method in (a). Reprinted with permission from [236], J. Li et al., *Appl. Phys. Lett.* 75, 367 (1999). © 1999, American Institute of Physics.

7.2.2. Horizontally Aligned Carbon Nanotubes

The first horizontally aligned growth of carbon nanotubes was successfully carried out by Kroto's group at the University of Sussex, UK, and Cheetham' group at the University of California at Santa Barbara, USA [245]. The cobalt film deposited on a silica substrate was etched into tracks 1–20 μm wide using a laser beam. Laser beam generates tracks free of cobalt, leaving cobalt particles evenly positioned along the edges of the tracks. Carbon nanotube growth was carried out using pyrolysis of 2-amino-4,6-dichloro-*s*-triazine. After growth, dense nanotubes were found lying inside tracks and aligned from one edge to the other edge of a track [245]. This is the first synthesis of horizontally aligned nanotubes. Dai et al. [63, 141, 203, 246–253] continued extensive research on controllable growth of horizontally aligned carbon nanotubes on patterned Si substrates. A network of suspended SWNTs was grown from Si towers [249]. These suspended SWNTs might be used for interconnect of nanodevices. In order to enhance the directionality of growth, the electric field is introduced horizontally, resulting in better aligned growth of SWNTs as shown in Figure 34 [252]. This was also inspired by the field-assisted growth of vertically aligned CNTs [254]. It was suggested that the self-directed growth between the neighboring pillars is due to the swing of the nanotube cantilever which contacts a catalyst particle in liquid phase as the nanotube grows [255]. These results confirm the possibility of self-assembled wiring of nanostructures. MWNTs were grown from the cleaved side of a $Si/SiO_2/Si$ mulitlayer structure to control the diameter of nanotubes [256]. In order for large-scale fabrication of nanodevices using CNTs, alternative selective placement of CNTs other than an AFM or STM is desirable. Controlled placement of SWNTs on prepatterned Si substrates has been achieved by a combination of the highly selective adsorption of SWNTs onto open regions of amino-functionalized SiO_2 in a polymer resist, followed by lift-off [257]. Controlled placement has also been achieved by electrical field-assisted alignment of CNTs under ac electric field [258]. Of particular interest is the utilization of nanotubes for nanolithography. One example is to use nanotubes as AFM tips to write nanopatterns [259]. Figure 35 shows the AFM image of a 10 nm wide line (SiO_2) fabricated by a nanotube tip [259]. Nanoelectrodes with a spacing of

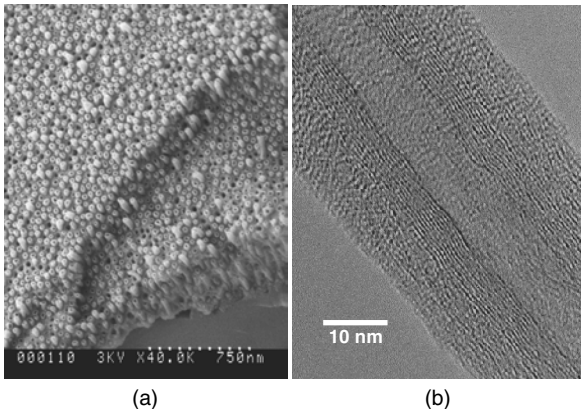


Figure 33. (a) Cross-sectional SEM images of CNT arrays (about 500 nm long) on Si substrate fabricated in our laboratories. (b) TEM image of a CNT grown from the Al_2O_3 template showing multiwalls. Adapted with permission from [243], W. C. Hu et al., *J. Nanosci. Nanotechnol.* 2, 203 (2002). © 2002, American Scientific Publishers.

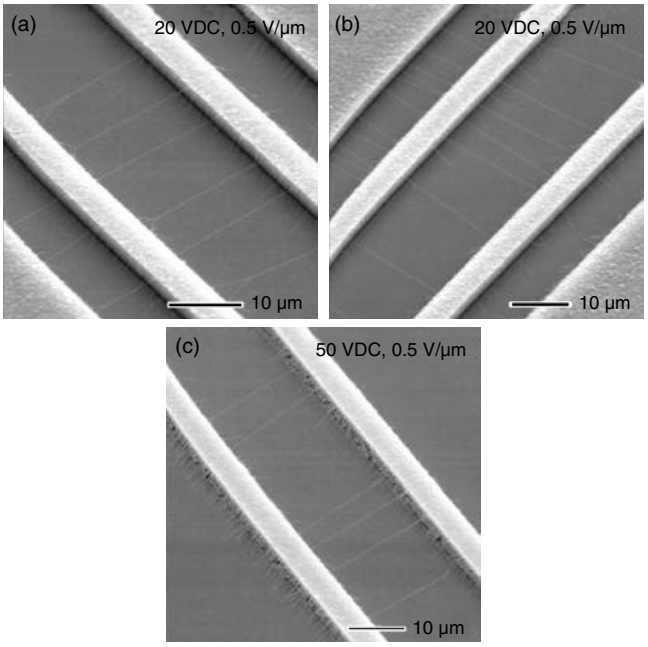


Figure 34. SEM images of suspended SWNTs grown in various electric fields. The spacing between the outer poly-Si electrodes is 40 μm in (a) and (b), and 100 μm in (c). Reprinted with permission from [252], Y. Zhang et al., *Appl. Phys. Lett.* 79, 3155 (2001). © 2001, American Institute of Physics.

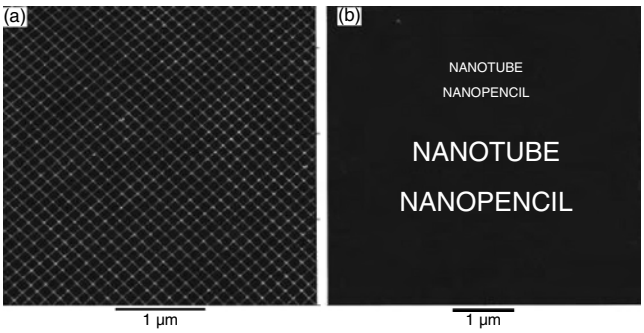


Figure 35. (a) AFM image of 2 nm tall, 10 mm wide, and 100 nm spaced silicon-oxide (light) lines fabricated by a nanotube tip. Bias voltage = −9 V; speed = 10 m/s; cantilever amplitude during lithography = 0.55 nm; cantilever amplitude = 9.6 nm. (b) Silicon oxide "words" written by the nanotube tip. Reprinted with permission from [259], H. Dai et al., *Appl. Phys. Lett.* 73, 1508 (1998). © 1998, American Institute of Physics.

~30 nm have been demonstrated using a carbon nanotube as a shadow mask, where a nanotube was suspended on PR supporters [260]. Figure 36 shows the schematic of the approach and Figure 37 shows the AFM image of the gold electrodes with a gap of ~30 nm [260]. Carbon nanotubes have also been used as masks for plasma etching where a ~20 nm nanowire beneath the nanotube was formed after etching [261].

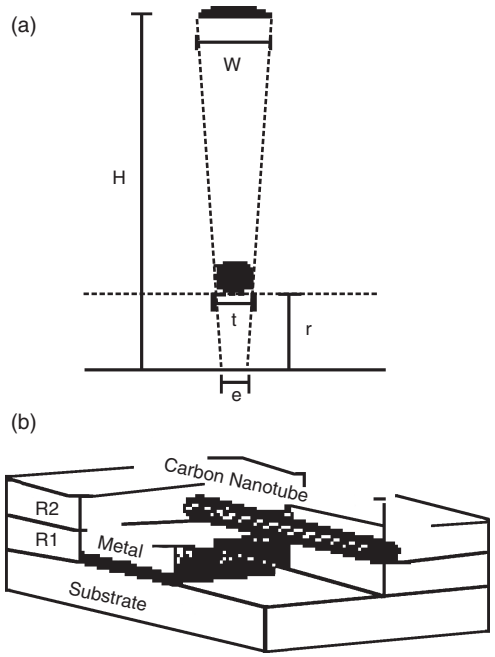


Figure 36. (a) In a shadow mask method, the shadow or gap size (ε) depends only on geometrical factors: the mask width or tube diameter (t), the tube–substrate distance (r), the evaporation source distance (H), and width (W). (b) Carbon nanotubes are incorporated between two layers of e-beam resist (R1 and R2) to act as a shadow mask preventing metal deposition at one point along a thin wire. The figure illustrates the process after metal deposition. Reprinted with permission from [260], J. Lefebvre et al., *Appl. Phys. Lett.* 76, 3828 (2000). © 2000, American Institute of Physics.

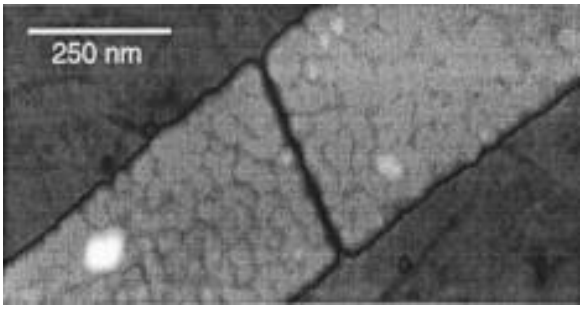


Figure 37. AFM image of a sub-30 nm gap in a 350 nm wide gold wire fabricated using a SWNT bundle. Individual SWNTs or small SWNT bundles (dark gray lines in the image) have been contacted using this method. Reprinted with permission from [260], J. Lefebvre et al., *Appl. Phys. Lett.* 76, 3828 (2000). © 2000, American Institute of Physics.

8. SUMMARY

Carbon nanotube-based nanoelectronics have been reviewed extensively, including electronic structure and transport properties of carbon nanotubes, various electronic devices, and nanotube-based nanofabrication. Since the discovery of CNTs for more than 10 years, great progress has been made in this field. Theoretical calculation predicted the one-dimensional energy band structure of single-wall carbon nanotubes with ballistic transport up to a few micrometers. Experiments successfully confirmed the theoretical predication. There have been demonstrations of the basic functions of carbon nanotube-based nanoelectronic devices such as single-electron transistors, field-effect transistors, logic gates, and memory devices. There is an almost limitless number of new geometries and topics waiting to be explored and all kinds of new structures to be created [262]. Upon comparison of their performance, carbon nanotube field-effect transistors are superior over Si MOS transistors. It is very promising to use carbon nanotube field-effect transistors as building blocks for large-scale integrated circuits when the current silicon technology reaches the end of its roadmap.

However, incorporation of the CNT devices to the existing Si manufacturing process is still far away. The current practice of fabrication of CNT transistors relies on arduous contacting of randomly distributed CNTs from suspension or time-consuming manipulation by AFM. Although great progress has been made in controllable growth and placement of carbon nanotubes, a planar process with high yield, which is the cornerstone of silicon technology, has not yet been achieved in the carbon nanotube process [263]. It is believed that in the next 10 to 20 years progress in process technologies for fabrication of carbon nanotube-based devices and circuits will eventually pave the way for electronics revolution in the molecular or nanoscale.

GLOSSARY

Atomic force microscope (AFM) A system with a mechanical tip scanning over a substrate to sense the atomic forces. The signal from the atomic forces can be synthesized into an image, which is capable to exhibit images of atoms of the substrate. It can also be used to manipulate atoms.

Carbon nanotube One or several graphite sheets that is wrapping together co-axially to form a tube-like material. If it has only on sheet, it is called single-wall carbon nanotube. If it has several sheets, it is called multi-wall carbon nanotube.

Field-effect transistor A three-terminal device with its two-terminal conductivity being totally controlled by the third terminal. In digital electronics, it usually functions as a switch.

Logic gate A device consisting field-effect transistors functions as a switch with "1" representing conduction and "0" representing non-conduction.

Memory devices A device that can store information in the form of "1" and "0."

Nanoelectronics Electronics built from nanoscale components.

Scanning tunneling microscope (STM) A vacuum system with an electrical tip scanning over a substrate to induce electrons tunneling into the substrate. The signal from the tunneling electrons can be synthesized into an image, which is capable to exhibit images of atoms of the substrate. It can also be used to manipulate atoms.

Single-electron transistors A device that exhibits staircase conductivity due to single-electron charging effect usually at low temperature.

ACKNOWLEDGMENT

This work was sponsored by National Science Foundation under the Materials Research Science and Engineering Center (MRSEC) program (DMR-9809686) and Nanoscale Exploratory Research (NER) program (ECS-0304129).

REFERENCES

1. S. Iijima, *Nature* 354, 56 (1991).
2. C. Dekker, *Phys. Today* 22 (May 1999).
3. P. M. Ajayan, *Chem. Rev.* 99, 1787 (1999).
4. M. S. Dresselhaus, G. Dresselhaus, and P. C. Eklund, "Science of Fullerenes and Carbon Nanotubes." Academic Press, San Diego, 1995.
5. R Saito, M. S. Dresselhaus, and G. Dresselhaus, "Physical Properties of Carbon Nanotubes." World Scientific, New York, 1998.
6. K. Tanaka, T. Yamabe, and K. Fukui, "The Science and Technology of Carbon Nanotubes." Elsevier, Oxford, 1999.
7. M. S. Dresselhaus, G. Dresselhaus, and Ph. Avouris, "Carbon Nanotubes." Springer-Verlag, Berlin, 2000.
8. J. W. Mintmire, B. I. Dunlap, and C. T. White, *Phys. Rev. Lett.* 68, 631 (1992).
9. N. Hamada, S.-I Sawada, and A. Oshiyama, *Phys. Rev. Lett.* 68, 1579 (1992).
10. R. Saito, M. Fujita, G. Dresselhaus, and M. S. Dresselhaus, *Appl. Phys. Lett.* 60, 2204 (1992).
11. S. Iijima and T. Ichibashi, *Nature* 363, 603 (1993).
12. D. S. Bethune, C. H. Kiang, M. S. DeVries, G. Gorman, R. Savoy, and R. Beyers, *Nature* 363, 605 (1993).
13. R. E. Smalley, http://www.ruf.rice.edu/~smalleyg/image_gallery.htm, 2002.
14. National Science and Technology Council, "National Nanotechnology Initiatrive: Leading to the Next Industrial Revolution," Washington, DC, 2000.
15. I. M. Ross, *Proc. IEEE* 86, 7 (1998).
16. J. S. Kilby, *IEEE Trans. Electron Dev.* ED-23, 648 (1976).
17. G. Binig, H. Roher, Ch. Gerber, and E. Weibel, *Phys. Rev. Lett.* 49, 57 (1982).
18. D. M. Eigler and E. K. Schweizer, *Nature* 344, 524 (1990).
19. M. F. Crommie, C. P. Lutz, and D. M. Eigler, *Science* 262, 218 (1993).
20. M. A. McCord and R. F. W. Pease, *J. Vac. Sci. Technol. B* 6, 293 (1988).
21. E. S. Snow, P. M. Campbell, and F. K. Perkins, *Proc. IEEE* 85, 601 (1997).
22. L. Stockman, G. Neuttiens, C. Van Haesendonck, and Y. Bruynseraede, *Appl. Phys. Lett.* 62, 2935 (1993).
23. A. L. DeLozanne, E. E. Ehrichs, and W. F. Smith, *J. Phys.: Condens. Matter.* 5, A409 (1993).
24. M. Wendel, S. Kuhn, H. Lorez, J. P. Kotthaus, and M. Holland, *Appl. Phys. Lett.* 65, 1775 (1994).
25. J. W. Lyding, T.-C. Shen, J. S. Hubacek, J. R. Tucker, and G. C. Abeln, *Appl. Phys. Lett.* 64, 2010 (1994).
26. J. E. Lilienfield, *U.S. Patent* 1, 745, 175, 1930.
27. O. Heil, *British Patent* 439457, 1935.
28. D. Kahng, *IEEE Trans. Electron Dev.* ED-23, 655 (1976).
29. P. S. Peercy, *Nature* 406, 1023 (2000).
30. J. D. Plummer and P. B. Griffin, *Proc. IEEE* 89, 240 (2001).
31. Semiconductor Industry Association (SIA), "International Technology Roadmap for Semiconductors," San Jose, CA, 2001.
32. Y. Taur, *IBM J. Res. Dev.* 46, 213 (2002).
33. H.-S. P. Wong, *IBM J. Res. Dev.* 46, 133 (2002).
34. D. J. Frank, R. H. Dennard, E. Nowark, P. M. Solomon, Y. Taur, and H.-S. P. Wong, *Proc. IEEE* 89, 259 (2001).
35. C. J. Gorter, *Physica* (*Amsterdam*) 17, 777 (1951).
36. T. A. Fulton and G. J. Dolan, *Phys. Rev. Lett.* 59, 109 (1987).
37. K. K. Likharev, *IBM J. Res. Dev.* 32, 144 (1988).
38. K. K. Likharev, *Proc. IEEE* 87, 606 (1999).
39. H. Ahmed, *J. Vac. Sci. Technol. B* 15, 2101 (1997).
40. D. Goldhaber-Gordon, M. S. Montemerlo, J. C. Love, G. J. Opiteck, and J. C. Ellenbogen, *Proc. IEEE* 85, 521 (1997).
41. W. Chen, H. Ahmed, and K. Nakazato, *Appl. Phys. Lett.* 66, 3383 (1995).
42. D. L. Klein, P. L. McEuen, J. E. B. Katari, R. Roth, and A. P. Alivisatos, *Appl. Phys. Lett.* 68, 2574 (1996).
43. Y. Nakamura, C. D. Chen, and J. S. Tsai, *Japan J. Appl. Phys.* 35, L1465 (1996).
44. J.-I. Shirakashi, K. Matzumoto, N. Miura, and M. Konagai, *Appl. Phys. Lett.* 72, 1893 (1998).
45. U. Meirav, M. A. Kastner, and S. J. Wind, *Phys. Rev. Lett.* 65, 771 (1990).
46. D. Ali and H. Ahmed, *Appl. Phys. Lett.* 64, 2119 (1994).
47. E. Leobandung, L. Guo, Y. Wang, and S. Y. Chou, *Appl. Phys. Lett.* 67, 938 (1995).
48. Y. Takahashi, M. Nagase, H. Namatsu, K. Kuihara, K. Iwadate, Y. Nakajima, S. Horiguchi, K. Murase, and M. Tabe, *Electron Lett.* 31, 136 (1995).
49. K. Kurihara, H. Namatsu, M. Nagase, and T. Takino, *Microelectron. Eng.* 35, 261 (1997).
50. L. Zhuang, L. Guo, and S. Y. Chou, *Appl. Phys. Lett.* 72, 1205 (1998).
51. L. Guo, E. Leobandung, and S. Y. Chou, *Science* 275, 649 (1997).
52. R. Smith and H. Ahmed, *J. Appl. Phys.* 81, 6 (1997).
53. K. Saito, F. A. Williams, and A. S. Gordon, *Combustion Sci. Technol.* 47, 117 (1986).
54. K. Saito, F. A. Williams, W. F. Stickle, and A. S. Gordon, in "Proceedings of 20th Fall Technical Meeting," Eastern Section, The Combustion Institute, Gaithersburg, MD, 1987.
55. K. Saito, private communication.
56. L. Yuan, K. Saito, C. Pan, F. A. Williams, and A. S. Gordon, *Chem. Phys. Lett.* 340, 237 (2001).

57. L. Yuan, K. Saito, W. Hu, and Z. Chen, *Chem. Phys. Lett.* 346, 23 (2001).
58. T. Guo, P. Nikolaev, A. Thess, D. T. Colbert, and R. E. Smalley. *Chem. Phys. Lett.* 243, 49 (1995).
59. T. Guo and R. E. Smalley, in "Recent Advances in the Chemistry and Physics of Fullerenes and Related Materials" (R. S. Ruoff and K. M. Kadish, Eds.), p. 626. The Electrochemical Society, Pennington, NJ, 1995.
60. A. Thess, R. Lee, P. Nikolaev, H. J. Dai, P. Petit, J. Robert, C. H. Xu, Y. H. Lee, S. G. Kim, A. G. Rinzler, D. T. Colbert, G. E. Scuseria, D. Tomanek, J. E. Fischer, and R. E. Smalley, *Science* 273, 483 (1996).
61. J. Kong, A. Cassell, and H. Dai, *Chem. Phys. Lett.* 292, 567 (1998).
62. A. Cassell, J. Raymakers, J. Kong, and H. Dai, *J. Phys. Chem.* 103, 6484 (1999).
63. J. Kong, H. T. Soh, A. Cassell, C. F. Quate, and H. Dai, *Nature* 395, 878 (1998).
64. R. Saito, M. Fujita, G. Dresselhaus, and M. S. Dresselhaus, *Phys. Rev. B* 46, 1804 (1992).
65. J. W. G. Wildoer, L. C. Venema, A. G. Rinzler, R. E. Smalley, and C. Dekker, *Nature* 391, 59 (1998).
66. M. S. Dresshaus, *Nature* 391, 19 (1998).
67. T. W. Odom, J.-L. Huang, P. Kim, and C. M. Lieber, *Nature* 391, 62 (1998).
68. J. A. Stroscio, R. M. Freenstra, and A. P. Fein, *Phys. Rev. Lett.* 57, 2579 (1986).
69. N. D. Lang, *Phys. Rev. B* 34, R5947 (1986).
70. A. M. Rao, P. C. Eklund, S. Bandow, A. Thess, and R. E. Smalley, *Nature* 388, 257 (1997).
71. A. M. Rao, E. Richter, S. Bandow, B. Chase, P. C. Eklund, K. A. Williams, S. Fang, K. R. Subbaswamy, M. Menon, A. Thess, R. E. Smalley, G. Dresselhaus, and M. S. Dresselhaus, *Science* 275, 187 (1997).
72. J. E. Fischer, H. Dai, A. Thess, R. Lee, N. M. Hanjani, D. L. Dehaas, and R. E. Smalley, *Phys. Rev. B* 55, R4921 (1997).
73. P. Petit, E. Jouguelet, J. E. Fischer, A. G. Rinzler, and R. E. Smalley, *Phys. Rev. B* 56, 9275 (1997).
74. T. Pichler, M. Knupfer, M. S. Golden, J. Fink, A. Rinzler, and R. E. Smalley, *Phys. Rev. Lett.* 80, 4729 (1998).
75. J.-C. Charlier and Ph. Lambin, *Phys. Rev. B* 57, R15037 (1998).
76. J.-C. Charlie and J.-P. Issi, *Appl. Phys. A* 67, 79 (1998).
77. M. J. O'Connell, S. M. Bachilo, C. B. Huffman, V. C. Moore, M. S. Strano, E. H. Haroz, K. L. Rialon, P. J. Boul, W. H. Noon, C. Kittrell, J. Ma, R. H. Hauge, R. B. Weisman, and R. E. Smalley, *Science* 297, 593 (2002).
78. P. Delaney, H. J. Choi, J. Lhm, S. G. Louie, and M. L. Cohen, *Nature* 391, 466 (1998).
79. S. G. Lemay, J. G. Janssen, M. van den Hout, M. Mooij, M. J. Bronlkowski, P. A. Willis, R. E. Smalley, L. P. Kouwenhoven, and C. Dekker, *Nature* 412, 617 (2001).
80. L. Langer, L. Stockman, J. P. Heremans, V. Bayot, C. H. Olk, C. Van Haesendonck, Y. Bruynseraede, and J.-P. Issi, *J. Mater. Res.* 9, 927 (1994).
81. W. A. de Heer, A. Chatelain, and D. Ugarte, *Science* 270, 1179 (1995).
82. L. Langer, V. Bayot, E. Grivei, J.-P. Issi, J. P. Heremans,, C. H. Olk, L. Stockman, C. Van Haesendonck, and Y. Bruynseraede, *Phys. Rev. Lett.* 76, 479 (1996).
83. H. Dai, E. W. Wong, and C. M. Lieber, *Science* 272, 523 (1996).
84. T. W. Ebbesen H. J. Lezec, H. F. Ghaemi, T. Thio, and P. A. Wolff, *Nature* 391, 667 (1996).
85. M. Bockrath, D. H. Cobden, P. L. McEuen, N. G. Chopra, A. Zettl, A. Thess, and R. E. Smalley, *Science* 275, 1922 (1997).
86. S. J. Tans, M. H. Devoret, H. Dai, A. Thess, R. E. Smalley, L. J. Geerligs, and C. Dekker, *Nature* 386, 474 (1997).
87. L. Chico, L. X. Benedict, S. G. Louie, and M. L. Cohen, *Phys. Rev. B* 54, 2600 (1996).
88. C. Roland, M. B. Nardelli, H. Guo, H. Mehrez, J. Taylor, J. Wang, and Y. Wei, *Surf. Rev. Lett.* 7, 637 (2000).
89. W. Tian and S. Datta, *Phys. Rev. B* 49, 5097 (1994).
90. M. F. Lin and K.-W.-K. Shung, *Phys. Rev. B* 51, 7592 (1995).
91. S. Frank, P. Poncharal, Z. L. Wang, and W. A. de Heer, *Science* 280, 1744 (1998).
92. J. Kong, E. Yenilmez, T. W. Tonbler, W. Kim, H. Dai, R. B. Laughlin, L. Liu, C. S. Jayanthi, and S. Y. Wu, *Phys. Rev. Lett.* 87, 106801 (2001).
93. C. T. White and T. N. Todorov, *Nature* 393, 240 (1998).
94. J. W. Park, J. Kim, J.-O. Lee, K.-C. Kang, J.-J. Kim, and K.-H. Yoo, *Appl. Phys. Lett.* 80, 133 (2002).
95. V. H. Crespi, A. M. L. Cohen, and A. Rubio, *Phys. Rev. Lett.* 79, 2093 (1997).
96. M. P. Anantram and T. R. Govindan, *Phys. Rev. B* 58, 4882 (1998).
97. T. Kostyrko, M. Bartkowiak, and G. D. Mahan, *Phys. Rev. B* 59, 3241 (1999).
98. T. Ando, T. Nakanishi, and R. Saito, *J. Phys. Soc. Jpn.* 67, 2857 (1998).
99. T. Ando, *Semicond. Sci. Technol.* 15, R13 (2000).
100. P. L. McEuen, M. Bockrath, D. Cobden, Y. G. Yoon, and S. Louie, *Phys. Rev. Lett.* 83, 5098 (1999).
101. S. Roche, F. Triozon, and A. Rubio, *Appl. Phys. Lett.* 79, 3690 (2001).
102. S. J. Tans, M. H. Devoret, R. J. A. Groeneveld, and C. Dekkers, *Nature* 397, 673 (1999).
103. D. H. Cobden, M. Bockrath, P. L. McEuen, A. G. Rinzler, and R. E. Smalley, *Phys. Rev. Lett.* 81, 681 (1998).
104. D. H. Cobden, M. Bockrath, N. G. Chopra, A. Zettle, P. L. McEuen, A. G. Rinzler, A. Thesis, and R. E. Smalley, *Physica B* 251, 132 (1998).
105. Y. Aharonove and D. Bohm, *Phys. Rev.* 115, 485 (1959).
106. W. Tian and S. Data, *Phys. Rev. B* 49, 5097 (1994).
107. A. Bachtold, C. Strunk, J.-P. Salvetat, J.-M. Bonard, L. Forro, T. Nussbaumer, and C. Schonenberger, *Nature* 397, 673 (1999).
108. S. J. Tans, M. H. Devoret, R. J. A. Groeneveld, and C. Dekker, *Nature* 394, 761 (1998).
109. M. P. A. Fisher and P. A. Glazman, "A Mesoscopic Electron Transport." Kluwer Academic, Boston, 1997.
110. C. Kane, L. Balents, and M. P. A. Fisher, *Phys. Rev. Lett.* 79, 5086 (1997).
111. R. Egger and A. O. Gogolin, *Phys. Rev. Lett.* 79, 5082 (1997).
112. R. Egger, *Phys. Rev. Lett.* 83, 5547 (1999).
113. M. Bockrath, D. H. Cobden, J. Lu, A. G. Rinzler, R. E. Smalley, L. Balents, and P. L. McEuen, *Nature* 397, 598 (1999).
114. J. Kong, C. Zhou, E. Yenilmez, and H. Dai, *Appl. Phys. Lett.* 77, 3977 (2000).
115. J. Kong, J. Cao, and H. Dai, *Appl. Phys. Lett.* 80, 73 (2002).
116. H. W. Ch. Postma, M. de Jonge, Z. Yao, and C. Dekker, *Phys. Rev. B* 62, R10653 (2000).
117. H. W. Ch. Postma, A. Sellmeijer, and C. Dekker, *Adv. Mater.* 17, 1299 (2000).
118. H. W. Ch. Postma, T. Teepen, Z. Yao, and C. Dekker, *Science* 293, 76 (2001).
119. J. B. Cui, M. Burghard, and K. Kern, *Nano Lett.* 2, 117 (2002).
120. J. W. Park, J. B. Choi, and K. H. Yoo, *Appl. Phys. Lett.* 81, 2644 (2002).
121. K. Ishibashi, M. Suzuki, T. Ida, and Y. Aoyagi, *Appl. Phys. Lett.* 79, 1864 (2001).
122. J. Nygard and D. H. Cobden, *Appl. Phys. Lett.* 79, 4216 (2001).
123. J. Park and P. L. McEuen, *Appl. Phys. Lett.* 79, 1363 (2001).
124. D. Bozovic and M. Bockrath, *Appl. Phys. Lett.* 78, 3693 (2001).
125. K. Ishibashi, M. Suzuki, T. Ida, and Y. Aoyagi, *Appl. Phys. Lett.* 80, 4238 (2002).
126. P. G. Collins, A. Zettle, H. Bando, A. Thess, and R. E. Smalley, *Science* 278, 100 (1997).

127. L. Roschier, J. Penttila, M. Martin, P. Hakonen, M. Paalanen, U. Tapper, E. I. Kauppinen, C. Journet, and P. Bernier, *Appl. Phys. Lett.* 75, 728 (1999).
128. L. Roschier, R. Tarkiainen, M. Ashlskog, M. Paalanen, and P. Hakonen, *Appl. Phys. Lett.* 78, 3295 (2001).
129. W. B. Choi, J. U. Chu, K. S. Jeong, E. J. Bae, and J.-W. Lee, *Appl. Phys. Lett.* 79, 3696 (2001).
130. M. Suzuki, K. Ishibashi, K. Toratani, D. Tsuya, and Y. Aoyagi, *Appl. Phys. Lett.* 81, 2273 (2002).
131. M. Ashlog, R. Tarkiainen, L. Roschier, and P. Hakonen, *Appl. Phys. Lett.* 77, 4307 (2000).
132. A. Kanda, Y. Ootuka, K. Tsukagoshi, and Y. Aoyagi, *Appl. Phys. Lett.* 79, 1354 (2001).
133. R. J. Schoelkopf, P. Wahlgren, A. A. Kozhevnikov, P. Selsing, and D. E. Prober, *Science* 280, 1238 (1998).
134. P. G. Collins, M. S. Fuhrer, and A. Zettle, *Appl. Phys. Lett.* 76, 894 (2000).
135. H. Ouacha, M. Willander, H. Y. Yu, Y. W. Park, M. S. Kabir, S. H. M. Persson, L. B. Kish, and A. Ouacha, *Appl. Phys. Lett.* 80, 1055 (2002).
136. S. J. Tans, A. R. M. Verschueren, and C. Dekker, *Nature* 393, 49 (1998).
137. R. Martel, T. Schmidt, H. R. Shea, T. Hertel, and Ph. Avouris, *Appl. Phys. Lett.* 73, 2447 (1998).
138. P. Avouris, T. Hertel, R. Martel, T. Schmidt, H. Shea, and R. Walkup, *Appl. Surf. Sci.* 141, 201 (1999).
139. P. Avouris, *Chem. Phys.* 281, 429 (2002).
140. P. G. Collins and P. Avouris, *Sci. Am.* 62 (2000).
141. H. Dai, J. Kong, C. Zhou, N. Franklin, T. Tombler, A. Cassell, S. Fan, and M. Chapline, *J. Phys. Chem. B* 103, 11246 (1999).
142. J. Guo, M. Lundstrom, and S. Datta, *Appl. Phys. Lett.* 80, 3192 (2002).
143. A. Rochefort, M. Di Ventra, and Ph. Avouris, *Appl. Phys. Lett.* 78, 2521 (2001).
144. S. J. Wind, J. Appenzeller, R. Martel, V. Derycke, and Ph. Avouris, *Appl. Phys. Lett.* 80, 3817 (2002).
145. B. Yu, in "Proc. Int. Electron Dev. Meet.," IEEE Electron Device Society, 2001, p. 937.
146. R. Chau, in "Proc. Int. Electron Dev. Meet.," IEEE Electron Device Society, 2001, p. 621.
147. R. S. Muller and T. I. Kamins, "Device Electronics for Integrated Circuits." Wiley, New York, 1986.
148. C. Bena, S. Vishveshwara, L. Balents, and M. P. A. Fisher, *Phys. Rev. Lett.* 89, 37901 (2002).
149. S. Rosenblat, Y. Yaish, J. Park, J. Gore, V. Sazonova, and P. L. McEuen, *Nano Lett.* 2, 869 (2002).
150. M. Kruger, M. R. Buitelaar, T. Nussbaumer, C. Schnenberger, and L. Forro, *Appl. Phys. Lett.* 78, 1291 (2001).
151. S. Kazaoui, N. Minami, N. Matsuda, H. H. Kataura, and Y. Achiba, *Appl. Phys. Lett.* 78, 3433 (2001).
152. M. S. Fuhrer, B. M. Kim, T. Durkop, and T. Brintlinger, *Nano Lett.* 2, 755 (2002).
153. K. Liu, M. Burghard, and S. Roth, *Appl. Phys. Lett.* 75, 2494 (1999).
154. T. W. Tombler, C. Zhou, J. Kong, and H. Dai, *Appl. Phys. Lett.* 76, 2412 (2000).
155. Y. H. Kim and K. J. Cheng, *Appl. Phys. Lett.* 81, 2264 (2002).
156. A. Javey, M. Shim, and H. Dai, *Appl. Phys. Lett.* 80, 1064 (2002).
157. C. Zhou, J. Kong, and H. Dai, *Appl. Phys. Lett.* 76, 1597 (2000).
158. J. O'Keeffe, C. Wei, and K. Cho, *Appl. Phys. Lett.* 80, 676 (2002).
159. J. Guo, S. Goasguen, M. Lundstrom, and S. Datta, *Appl. Phys. Lett.* 81, 1486 (2002).
160. A. Bachtold, P. Hadley, T. Nakanishi, and C. Dekker, *Science* 294, 1317 (2001).
161. V. Derycke, R. Martel, J. Appenzeller, and Ph. Avouris, *Nano Lett.* 1, 453 (2001).
162. X. Liu, C. Lee, and C. Zhou, *Appl. Phys. Lett.* 79, 3329 (2001).
163. T. Rueckes, K. Kim, E. Joselevich, G. Y. Tseng, C.-L. Cheung, and C. M. Lieber, *Science* 289, 94 (2000).
164. M. Radosavljevic, M. Fretag, K. V. Thadani, and A. T. Johson, *Nano Lett.* 2, 761 (2002).
165. J. B. Cui, R. Sordan, M. Burghard, and K. Kern, *Appl. Phys. Lett.* 81, 3260 (2002).
166. N. Yoneya, K. Tsukagoshi, and Y. Aoyagi, *Appl. Phys. Lett.* 81, 2250 (2002).
167. R. D. Antonov and A. T. Johnson, *Phys. Rev. Lett.* 83, 3274 (1999).
168. R. S. Lee, H. J. Kim, J. E. Fischer, J. Lefebvre, M. Radosavljevic, J. Hone, and A. T. Johnson, *Phys. Rev. B* 61, 4526 (2000).
169. M. Bockrath, J. Hone, A. Zettle, P. L. McEuen, A. G. Rinzler, and R. E. Smalley, *Phys. Rev. B* 61, R10606 (2000).
170. J. Kong, C. Zhou, E. Yenilmez, and H. Dai, *Appl. Phys. Lett.* 77, 3977 (2000).
171. V. Derycke, R. Martel, J. Appenzeller, and Ph. Avouris, *Appl. Phys. Lett.* 80, 2773 (2002).
172. Z. Yao, H. W. Ch. Postma, L. Balents, and C. Dekker, *Nature* 402, 273 (1999).
173. A. N. Andriotis, M. Menon, D. Srivastava, and L. Chernozatonskii, *Phys. Rev. Lett.* 87, 66802 (2001).
174. B. Wei, R. Spolenak, P. Kohler-Redlich, M. Ruhle, and E. Arzt, *Appl. Phys. Lett.* 74, 3149 (1999).
175. R. Pati, Y. Zhang, and S. K. Nayak, *Appl. Phys. Lett.* 81, 2638 (2002).
176. C. Zhou, J. Kong, E. Yenilmez, and H. Dai, *Science* 290, 1552 (2000).
177. L. Chico, V. H. Crespi, L. X. Benedict, S. G. Louie, and M. L. Cohen, *Phys. Rev. Lett.* 76, 971 (1996).
178. J. Han, M. P. Anantram, R. L. Jaffe, J. Kong, and H. Dai, *Phys. Rev. B* 57, 14983 (1998).
179. J.-O. Lee, H. Oh, J.-R. Kim, K. Kang, J.-J. Kim, J. Kim, and K.-H. Yoo, *Appl. Phys. Lett.* 79, 1351 (2001).
180. G. W. Ho, A. T. S. Wee, and J. Lin, *Appl. Phys. Lett.* 79, 260 (2001).
181. X. Ma and E. G. Wang, *Appl. Phys. Lett.* 78, 978 (2001).
182. M. S. Fuhrer, J. Nygard, L. Shih, M. Forero, Y.-G. Yoon, M. S. C. Mazzoni, H. J. Choi, J. Ihm, S. G. Loie, A. Zettle, and P. L. McEuen, *Science* 288, 494 (2000).
183. F. Leonard and J. Tersoff, *Phys. Rev. Lett.* 85, 4767 (2000).
184. T. Yamada, *Appl. Phys. Lett.* 80, 4027 (2002).
185. A. A. Odintsov, *Phys. Rev. Lett.* 85, 150 (2000).
186. G. D. J. Smit, S. Rogge, and T. M. Klapwijk, *Appl. Phys. Lett.* 81, 3852 (2002).
187. M. Menon and D. S. Srivastava, *Phys. Rev. Lett.* 79, 4453 (1997).
188. J. Li, C. Papadopoulos, and J. Xu, *Nature* 402, 254 (1999).
189. C. Papadopoulos, A. Rakitin, J. Li, A. S. Vedeneev, and J. M. Xu, *Phys. Rev. Lett.* 85, 3476 (2000).
190. B. C. Satishkumar, P. J. Thomas, A. Govindaraj, and C. N. R. Rao, *Appl. Phys. Lett.* 77, 2530 (2000).
191. D. Zhou and S. Seraphin, *Chem. Phys. Lett.* 238, 286 (1995).
192. B. Gan, J. Ahn, A. Zhang, S. F. Yoon, Q. F. Huang, H. Yang, M. B. Yu, and W. Z. Li, *Diamond Relat. Mater.* 9, 897 (2000).
193. J. M. Ting and C.-C. Chang, *Appl. Phys. Lett.* 80, 324 (2002).
194. B. Gan, J. Ahn, Q. Zhang, Q.-F. Huang, C. Kerlit, S. F. Yoon, Rusli, V. A. Ligachev, X.-B. Zhang, and W.-Z. Li, *Mater. Lett.* 45, 315 (2000).
195. W. Z. Li, J. G. Wen, and Z. F. Ren, *Appl. Phys. Lett.* 79, 1879 (2001).
196. P. Nagy, R. Ehlich, L. P. Biro, and J. Gyulai, *Appl. Phys. A* 70, 481 (2000).
197. A. N. Andriotis, M. Menon, D. Srivastva, and L. Chernozatonskii, *Phys. Rev. Lett.* 87, 66802 (2001).
198. A. N. Andriotis, M. Menon, D. Srivastva, and L. Chernozatonskii, *Appl. Phys. Lett.* 79, 266 (2001).
199. M. Menon and D. Srivastva, *J. Mater. Res.* 13, 2357 (1998).
200. M. Freitag, M. Radosavljevic, Y. Zhou, and A. T. John, *Appl. Phys. Lett.* 79, 3326 (2001)

201. Y. Zhang, T. Ichihashi, E. Landree, F. Nihey, and S. Iijima, *Science* 285, 1719 (1999).
202. J. Appenzeller, R. Martel, P. Avouris, H. Stahl, and B. Lengeler, *Appl. Phys. Lett.* 78, 3313 (2001).
203. H. T. Soh, C. F. Quate, A. F. Morpurgo, C. M. Marcus, J. Kong, and H. Dai, *Appl. Phys. Lett.* 75, 627 (1999).
204. J. Lefebvre, J. F. Lynch, M. Llaguno, M. Radosavljevic, and A. T. Johnson, *Appl. Phys. Lett.* 75, 3014 (1999).
205. P. G. Collins, M. S. Arnold, and P. Avouris, *Science* 292, 706 (2000).
206. T. Hertel, R. Martel, and P. Avouris, *J. Phys. Chem. B* 102, 910 (1998).
207. L. C. Venema, J. W. G. Wildoer, H. L. J. T. Tuinstra, C. Dekker, A. G. Rinzler, and R. E. Smalley, *Appl. Phys. Lett.* 71, 2629 (1997).
208. A. Bachtold, M. S. Fuhrer, S. Plyasunov, M. Forero, E. H. Anderson, A. Zettle, and P. L. McEuen, *Phys. Rev. Lett.* 84, 6082 (2000).
209. M. Bockrath, W. Liang, D. Bozovic, J. H. Hafner, C. M. Lieber, M. Tinkham, and H. Park, *Science* 291, 283 (2001).
210. S. J. Tans and C. Dekker, *Nature* 404, 834 (2000).
211. M. R. Falvo, R. M. Taylor II, A. Helser, V. Chi, F. P. Brooks, Jr., S. Washburn, and R. Superfine, *Nature* 397, 236 (1997).
212. M. R. Falvo, G. J. Clary, R. M. Taylor II, V. Chi, F. P. Brooks, Jr., S. Washburn, and R. Superfine, *Nature* 389, 582 (1997).
213. D. Bozovic, M. Bockath, J. H. Halfner, C. M. Lieber, H. Park, and M. Tinkam, *Appl. Phys. Lett.* 78, 3693 (2001).
214. C. Thelander, M. H. Magnusson, K. Deppert, L. Samuelson, P. R. Poulsen, J. Nygard, and J. Borgreen, *Appl. Phys. Lett.* 78, 3693 (2001).
215. P. A. Williams, S. J. Padakis, M. R. Falvo, A. M. Patel, M. Sinclair, A. Seeger, A. Helser, R. M. Taylor II, S. Washburn, and R. Superfine, *Appl. Phys. Lett.* 80, 2574 (2002).
216. J.-Y. Park, Y. Taish, M. Brink, S. Rosenblatt, and P. L. McEuen, *Appl. Phys. Lett.* 80, 4446 (2002).
217. T. W. Thombler, C. Zhou, L. Alexseyev, J. Kong, H. Dai, L. Liu, C. S. Jayanthi, M. Tang, and S.-Y. Wu, *Nature* 405, 769 (2000).
218. L. Liu, C. S. Jayanthi, M. Tang, S. Y. Wu, T. W. Tombler, C. Zhou, L. Alexseyev, J. Kong, and H. Dai, *Phys. Rev. Lett.* 84, 4950 (2000).
219. L. Liu, C. S. Jayanthi, and S. Y. Wu, *Phys. Rev. B* 64, 33412 (2001).
220. H. Watanabe, C. Manabe, T. Shigematsu, and M. Shimizu, *Appl. Phys. Lett.* 78, 2928 (2001).
221. H. Hirayama, Y. Kawamoto, Y. Ohshima, and K. Takayanagi, *Appl. Phys. Lett.* 79, 1169 (2001).
222. P. J. de Pablo, C. Gomez-Navarro, A. Gil, J. Colchero, M. T. Martinez, A. M. Benito, W. K. Maser, J. Gomez-Herrero, and A. M. Baro, *Appl. Phys. Lett.* 79, 2979 (2001).
223. S. Akita, Y. Nakayama, S. Mizooka, Y. Takano, T. Okawa, Y. Miyatake, S. Yamanaka, M. Tsuji, and T. Nosaka, *Appl. Phys. Lett.* 79, 1691 (2001).
224. B. Q. Wei, R. Vajtai, and P. M. Ajayan, *Appl. Phys. Lett.* 79, 1172 (2001).
225. W. Z. Li, S. S. Xie, L. X. Qian, B. H. Chang, B. S. Zou, W. Y. Zhou, R. A. Zhao, and G. Wang, *Science* 274, 1701 (1996).
226. W. Z. Li, H. Zhang, C. Wang, Y. Zhang, L. Xu, K. Zhu, and S. Xie, *Appl. Phys. Lett.* 70, 2684 (1997).
227. D. Xu, G. Guo, L. Gui, Y. Tang, Z. Shi, Z. Jin, Z. Gu, W. Liu, X. Li, and G. Zhang, *Appl. Phys. Lett.* 75, 481 (1999).
228. Z. F. Ren, Z. P. Huang, J. W. Xu, J. H. Wang, P. Bush, M. P. Siegal, and P. N. Provencio, *Science* 282, 1105 (1998).
229. Z. F. Ren, Z. P. Huang, D. Z. Wang, J. G. Wen, J. W. Xu, J. H. Wang, L. E. Calvet, J. Chen, J. F. Klemic, and M. A. Reed, *Appl. Phys. Lett.* 75, 1086 (1999).
230. M. P. Siegal, D. L. Overmyer, and P. P. Provencio, *Appl. Phys. Lett.* 80, 2171 (2001).
231. S. Fan, W. Liang, H. Dang, N. Franklin, T. Tombler, M. Chapline, and H. Dai, *Physica E* 8, 179 (2000).
232. G. C. Wood and J. P. O'Sullivan, *J. Electrochem. Soc.* 116, 1351 (1969).
233. O. Jessensky, F. Muller, and U. Gosele, *J. Electrochem. Soc.* 145, 3735 (1998).
234. H. Masuda and K. Fukuda, *Science* 268, 1466 (1995).
235. J. Li, M. Moskovits, and T. L. Haslett, *Chem. Mater.* 10, 1963 (1998).
236. J. Li, C. Papadopoulos, J. M. Xu, and M. Moskovits, *Appl. Phys. Lett.* 75, 367 (1999).
237. J. S. Suh and J. S. Lee, *Appl. Phys. Lett.* 75, 2047 (1999).
238. J. S. Suh, K. S. Jeong, J. S. Lee, and I. Han, *Appl. Phys. Lett.* 80, 2392 (2002).
239. S. L. Sung, S. H. Tsai, C. H. Tseng, F. K. Chiang, X. W. Liu, and H. C. Shih, *Appl. Phys. Lett.* 74, 197 (1999).
240. S.-H. Jeong, O.-J. Lee, K.-H. Lee, S. H. Oh, and C.-G. Park, *Chem. Mater.* 14, 1859 (2002).
241. Z. H. Yuan, H. Huang, H. Y. Dang, J. E. Cao, B. H. Hu, and S. S. Fan, *Appl. Phys. Lett.* 78, 3127 (2001).
242. T. Iwasaki, T. Motoi, and T. Den, *Appl. Phys. Lett.* 75, 2044 (1999).
243. W. C. Hu, L. M. Yuan, Z. Chen, D. W. Gong, and K. Saito, *J. Nanosci. Nanotechnol.* 2, 203 (2002).
244. W. C. Hu, D. W. Gong, Z. Chen, L. M. Yuan, K. Saito, P. Kichambare, and C. A. Grimes, *Appl. Phys. Lett.* 79, 3083 (2001).
245. M. Terrones, N. Grobert, J. Olivares, J. P. Zhang, H. Terrones, K. Kordaros, W. K. Hsu, J. P. Hare, P. D. Towbsend, K. Prassides, A. K. Cheetham, H. W. Kroto, and D. R. M. Walto, *Nature* 388, 52 (1997).
246. H. Dai, *Phys. World* 43 (2000).
247. H. Dai, J. Kong, C. Zhou, N. Franklin, T. Tombler, A. Cassell, S. Fan, and M. Chapline, *J. Phys. Chem.* 103s, 11246 (1999).
248. A. M. Cassell, R. Franklin, T. W. Tombler, E. M. Chan, J. Han, and H. Dai, *J. Am. Chem. Soc.* 121, 7975 (1999).
249. N. R. Franklin and H. Dai, *Adv. Mater.* 12, 890 (2000).
250. N. R. Franklin, Y. Li, R. J. Chen, A. Javey, and H. Dai, *Appl. Phys. Lett.* 79, 4571 (2001).
251. N. R. Franklin, Q. Wang, T. W. Tombler, A. Javey, and M. Shim, *Appl. Phys. Lett.* 81, 913 (2002).
252. Y. Zhang, A. Chang, J. Cao, Q. Wang, W. Kim, Y. Li, N. Morris, E. Yenilmez, J. Kong, and H. Dai, *Appl. Phys. Lett.* 79, 3155 (2001).
253. A. Ural, Y. Li, and H. Dai, *Appl. Phys. Lett.* 81, 3464 (2002).
254. Y. Avigal and R. Kalish, *Appl. Phys. Lett.* 78, 2291 (2001).
255. Y. Homma, Y. Kobayashi, and T. Ogino, *Appl. Phys. Lett.* 81, 2261 (2002).
256. N. Chopra, P. D. Kichamare, R. Andrews, and B. J. Hinds, *Nano Lett.* 2, 1177 (2002).
257. J. C. Lewenstein, T. P. Burgin, A. Ribayrol, L. A. Nagahara, and R. K. Tsui, *Nano Lett.* 2, 443 (2002).
258. X. Q. Chen, T. Saito, H. Yamada, and K. Matsushige, *Appl. Phys. Lett.* 78, 3714 (2001).
259. H. Dai, N. Franklin, and J. Han, *Appl. Phys. Lett.* 73, 1508 (1998).
260. J. Lefebvre, M. Radosavljevic, and A. T. Johnson, *Appl. Phys. Lett.* 76, 3828 (2000).
261. W. S. Yun, J. Kim, K.-H. Park, J. S. Ha, Y.-J. Ko, K. Park, S. K. Kim, Y.-J. Doh, H.-J. Lee, J. P. Salvetat, and L. Forro, *J. Vac. Sci. Technol. A* 18, 1329 (2000).
262. P. L. McEuen, *Phys. World* 31 (2000).
263. F. Kreupl, A. Graham, and W. Honlein, *Solid State Technol.* S9 (2002).